18 LABORATORY QUALITY CONTROL

18.1 Introduction

This chapter addresses internal laboratory quality control (QC), the purpose of which is to monitor performance, identify problems, and initiate corrective action. If project requests are more stringent than typical laboratory QC needs, the project manager and the laboratory should confer to see whether the laboratory can accommodate the tightened QC requirements. Laboratory data should be produced in a quality system[1] that incorporates planning, implementing, and internal assessment of the work performed by the laboratory, including QC. While this chapter focuses on laboratory QC, MARLAP fully endorses the need for a laboratory quality system and a Quality Manual that delineates the quality assurance (QA) policies and QC practices of the laboratory. General requirements for testing laboratories can be found in ISO/IEC 17025.

The chapter's purpose is to provide guidance to laboratory staff on those activities and professional practices a radioanalytical laboratory should undertake to produce data of known quality. This chapter also shows how to use statistical techniques to monitor specific measures of the analytical process to indicate the level of control of the analytical process within the laboratory. These measures are called "performance indicators," and the statistical techniques involve the use of control charts. Monitoring performance indicators through control charts enables the identification of trends. The laboratory can then address analytical problems and help improve the analytical process. Section 18.3.2 and Attachment 18A at the end of this chapter provide examples of several types of charts. The use of statistical techniques is the preferred method for implementing quality control in the laboratory (Attachment 18B). The chapter also identifies specific performance indicators, the principles that govern their use, indications and underlying causes of excursions, statistical means of evaluating performance indicators, and examples of root-cause evaluations.

The control of the analytical process in the laboratory is distinct from meeting the typical analytical needs of a specific project. This chapter addresses the former, to the extent that QC provides quantitative estimates of analysis and measurement controls that can be used to determine compliance with project objectives.

---

[1] A quality system is a structured and documented management framework that describes the policies, objectives, principles, organizational authority, responsibilities, accountability, and implementation plan of an organization for ensuring quality in its work processes, products (items), and services. The quality system provides for planning, implementing, and assessing the work performed by the organization and for carrying out required quality assurance and quality control (ANSI/ASQC E4, 1994).

18.1.1 Organization of Chapter

Chapter 18 has five major sections in addition to this introduction. Section 18.2 provides a general overview of QC and its application in the laboratory setting. Section 18.3 discusses the importance of evaluating performance indicators and provides statistical means for their evaluation. Sections 18.4 and 18.5 identify primary radiochemistry and instrumentation performance indicators, respectively, and discuss each in detail. Section 18.6 discusses other aspects of the analytical process that require scrutiny but are not formally considered performance indicators.

18.1.2 Format

The chapter is presented in a different format than the preceding chapters in order to highlight the performance indicators and to give examples. For each performance indicator, general guidance is provided in the format shown below.

Issue: Defines and summarizes the performance indicator

Discussion: Identifies those matters important to the performance indicator, including:

- What is the performance indicator and how does it work?
- Why is the performance indicator important, and what is its impact on the quality of the measurement?
- What is the relationship of the performance indicator and the combined standard uncertainty derived for the analytical method?
- What are the acceptable limits of the performance indicator?
- What are the key assumptions underlying the performance indicator?
- What limits and cautions are associated with the assumptions made?
- How sensitive is the quality of the measurement to the assumptions made?
- What is the appropriate frequency for assessing this performance indicator?

Excursions: "Excursions" are departures from the expected condition. This section addresses the

likely types of excursions encountered during laboratory analysis and explains what each may indicate. This section also discusses the potential reasons for these excursions and the implications for the analytical results.

Examples: Where appropriate, this section provides typical examples of excursions, potential reasons for excursions, and additional information.

18.2 Quality Control

Quality control includes all technical activities that measure the attributes and performance of a process, item, or service against defined standards to verify that they meet the stated requirements established by the customer. It also includes operational techniques and activities that are used to fulfill requirements for quality (ANSI/ASQC E4, 1994).

QC may not always detect blunders. Good laboratory practices, in addition to adherence to standard operating procedures (SOPs), are part of the overall QA/QC aspects needed to check the laboratory's performance. To monitor and control quality, laboratories use performance indicators, which are instrument- or protocol-related parameters that are routinely monitored to assess the laboratory's estimate of measurement uncertainty, precision, bias, etc. Initially, these parameters are used to maintain or demonstrate control over the analytical process. The performance indicators should be tracked by appropriate personnel. If the performance indicator control limits are exceeded, management should be informed and corrective action should be initiated.

Table 18.1 lists some of the potential causes for radioanalytical control excursions. By no means is the list complete, and the reader should be aware of additional potential causes of excursions that are presented in the rest of this chapter and the other chapters. Many problems are complex and have multiple components that could complicate the search for causes of protocol or instrument related excursions. A metrologist or radiochemist should be consulted to identify and remedy any analytical problems.

TABLE 18.1 — Problems leading to loss of analytical control

Radiochemical Processing	Source Preparation	Instrument Related	Other
Laboratory blunder	Laboratory blunder	Laboratory blunder	Laboratory blunder
Processing difficulty	Poor mounting	Electronic malfunction • preamplifier • power supply	Data transcription error
	Poor plating	• guard	

Radiochemical Processing	Source Preparation	Instrument Related	Other
Questionable reagent purity	Improper geometry	• analog to digital convertor (ADC) • gain • high voltage • discriminator • pole zero • shape constant	Incorrect units
Low tracer/carrier recovery	Incorrect thin plastic film thickness		Calculation error
Excessive tracer/carrier recovery	Improper plating on the planchet	Improper source or sample geometry	Software limitation
Inaccurate aliquanting of tracer/carrier	Excessive source mass	Poor counting statistics	Computer problem
Sample aliquanting inaccuracy	Uncorrected self absorption	Poor detector resolution	Loss of electrical power
Cross-contamination	Quenching	Detector contamination	Electrical power fluctuations
Inadequate dissolution of sample	Recoil contamination	Inappropriate/out-of-date efficiency, background or calibration factor	Mislabeling
Complex matrix		Background shift	Loss of sample
Sample heterogenity		Incorrect nuclear transformation data or other constants	Insufficient sample information
		Variable memory effects	
		Peak/calibration shift	Data processing problem
		Counting gas • pressure too high, too low, or variable • gas impurity	Interfering radionuclides
		Loss of vacuum/coolant	
		Temperature and humidity fluctuation	
		Measurement problem	

18.3 Evaluation of Performance Indicators

18.3.1 Importance of Evaluating Performance Indicators

As stated previously, performance indicators are measures of the analytical process that the laboratory monitors as part of its routine QC program. Performance indicators demonstrate whether the analytical process is performing as planned, when it has exhibited a statistical anomaly that requires investigation, and when a system has failed. Accordingly, monitoring performance indicators using established statistical techniques provides the laboratory with an effective tool for self assessment that allows the identification of trends or conditions that, while still within the established bounds of acceptability, are drifting or trending out of control. These

conditions can be addressed prospectively, allowing the laboratory to maintain analytical control. Additionally, this process allows the development of a data base regarding a protocol's or system's behavior over time or under a specified set of conditions.

18.3.2 Statistical Means of Evaluating Performance Indicators — Control Charts

The primary tool for statistical quality control is the control chart (see Attachment 18A). The theory that underlies a control chart is statistical hypothesis testing (see Appendix C). The implementation of a control chart makes the theory transparent to the average user and reduces the process of statistical inference to answering simple questions, such as, "Is the measured parameter greater than the upper control limit?" or "Is the measured parameter in the warning region?"

In theory, to test whether a parameter θ is above or below a certain value θ_0, a test statistic is defined and its distribution is determined under the assumption that $\theta = \theta_0$ (the null hypothesis). The value of the statistic is calculated and compared to critical values to test the assumption. In practice, a control chart is designed so that a non-statistician can perform these tests easily by comparing the measured value of the parameter to control limits and warning limits.

Most control charts do not implement hypothesis tests in a rigorous manner that allows decision error rates to be precisely determined. The charts are intended to be simple and practical tools for use even in situations where the assumptions needed for a rigorous test are not verifiable.

Every control chart has control limits, which define the acceptable range of the monitored variable. Many charts have both upper and lower limits. However, when changes in only one direction are of concern, only one limit is necessary. Most control charts have a central line, or reference line, which is an estimate of the expected value of the monitored variable. Many control charts also have warning limits, which lie between the central line and the control limits.

By definition, control limits are action limits. A single measured value that falls outside these limits requires that one stop the measurement process, investigate the problem, and if necessary take corrective action. The warning limits are optional but recommended, since they help one to identify and investigate possible problems before control limits are exceeded.

Types of Control Charts: Control charts based on grouped observations often are more powerful tools for detecting shifts of the monitored variable than charts based on individual observations. *Average charts*, or $\overline{X}$ *charts*, are used to monitor the arithmetic means of measured values obtained in "rational subgroups," which are subgroups of equal size chosen to ensure that the

measurement variability within each subgroup is likely to represent only the inherent variability of the measurement process produced by non-assignable causes (see Attachment 18A). When an $\overline{X}$ chart is used, a *range chart*, or *R chart*, is generally used in tandem to monitor within-group variability. (The *range* of a set of values is the difference between the largest value and the smallest.)

A control chart for individual values (*X chart* or *I chart*) is used when it is impractical to obtain measured values in the groups needed for an $\overline{X}$ chart. In this case, a *moving range chart* (*MR chart*) is often used as well to monitor variability. The moving range chart is an *R* chart based on the absolute differences between consecutive measured values.

A control chart may or may not be based on a particular type of data distribution. Most control charts use limits derived from the normal distribution but are intended to be used for data with almost any distribution (ISO 8258). However, when data obtained from radiation counters are monitored, the Poisson distribution may often be assumed. The standard types of control charts for Poisson data in industrial applications are called "*c* charts" (for total counts) and "*u* charts" (for count rates). A third type of Poisson control chart, which is a variant of the *u* chart, is frequently used to monitor radiation counter efficiency. When the data distribution is Poisson, separate charts for monitoring the value of the parameter and its variability are generally unnecessary because the mean and variance of a Poisson distribution are equal.

The following documents provide more guidance on the use of control charts:

- ASTM D6299. *Standard Practice for Applying Statistical Quality Assurance Techniques to Evaluate Analytical Measurement System Performance*.
- ASTM E882. *Standard Guide for Accountability and Quality Control in the Chemical Analysis Laboratory*. ANSI/ISO/ASQC A3534-2. *Statistics–Vocabulary and Symbols–Statistical Quality Control*.
- ISO 7870. *Control Charts – General Guide and Introduction*.
- ISO 7873. *Control Charts for Arithmetic Average with Warning Limits*.
- ISO 7966. *Acceptance Control Charts*.
- ISO 8258. *Shewhart Control Charts*.

- American Society for Testing and Materials (ASTM) MNL 7, *Manual on Presentation of Data and Control Chart Analysis* ASTM Manual Series, 6th Edition, 1990.

Figure 18.1 illustrates a typical control chart using counting data of a standard reference material (with limits corrected for decay) showing the statistical nature of the chart. The applicability of control chart techniques is based on the assumption that laboratory data approximate a normal distribution like that shown on the left of the vertical axis in the figure. The counting data plotted graphically represent the test results on the vertical axis and the scale order or time sequence in which the measurements were obtained on the horizontal axis. The mean of the measurements is represented by the central line (CL), and the limits of dispersion in terms of standard deviation are represented by the upper and lower warning and control limits (UWL, UCL, LWL, LCL). The warning limits are usually 2 standard deviations from the mean and the control limits are 3 standard deviations from the mean.

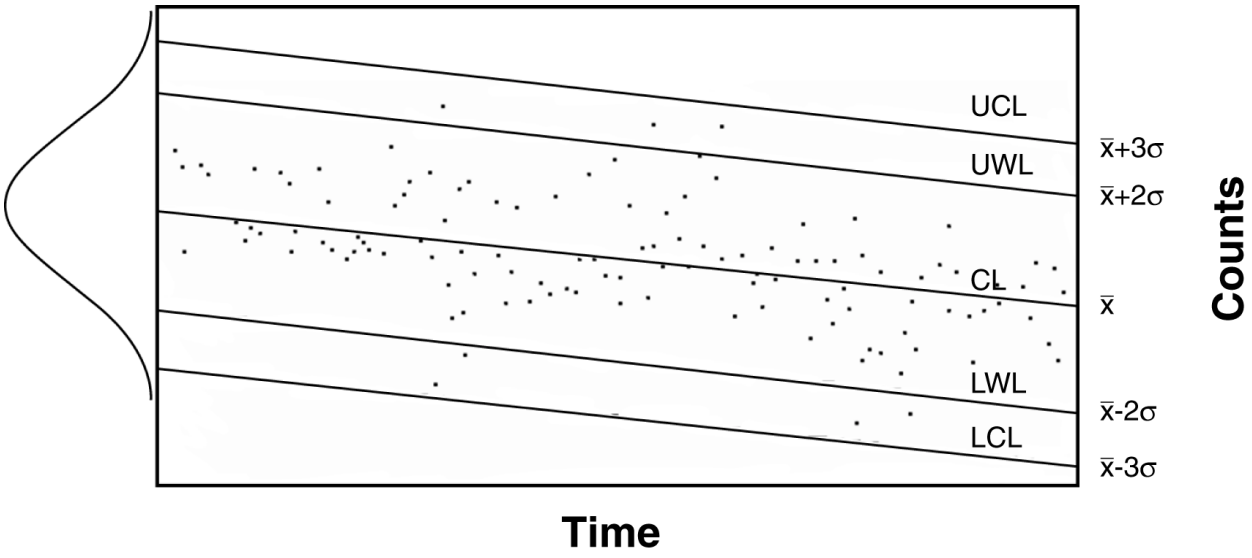

FIGURE 18.1 — Control chart for daily counting of a standard reference source, with limits corrected for decay. Statistical nature of chart is illustrated on the left by the Gaussian curve.

18.3.3 Measurement Uncertainty

Issue: Since laboratory radioactivity measurements always involve uncertainty, every measured result is uncertain to some degree. If the measurement uncertainties are large relative to the tolerances needed for decision making, the data may not be useful for their intended purpose. A discussion of measurement uncertainty is contained in Chapter 19, and the terms used in this section are defined in that chapter and in the Glossary.

Discussion: In order to determine the significance of a sample result, all reported values should be accompanied by the laboratory's best estimate of the uncertainty associated with the result. The "combined standard uncertainty" (one-sigma uncertainty) is obtained by propagating the uncertainties of all the input quantities that contribute to the calculation of the derived value (Chapter 19).

The combined standard uncertainty is used to indicate the statistical confidence in interpreting the performance indicator's ability to assess analytical quality. The estimated statistical confidence level that is usually associated with 1 combined standard uncertainty is about 68 percent, the confidence level for 2 combined standard uncertainties is about 95 percent, and the confidence level for 3 combined standard uncertainties is about 99 percent. It is important that the combined standard uncertainty be a fair estimate because it will indicate when the analytical process could be approaching the limits of statistical control and corrective actions should be initiated. A performance indicator exceeding ±2 combined standard uncertainty limits from the indicator's historical mean value may indicate that corrective action should be considered, and a performance indicator exceeding ±3 combined standard uncertainty limits from the indicator's historical mean value may indicate that an investigation must be conducted and corrective action may be necessary. Because statistical confidence never reaches 100 percent, it probably would be prudent to confirm the measurement for the performance indicator when it exceeds ±2 combined standard uncertainty limits. If the performance indicator value for repeat measurements do not exceed ±2 combined standard uncertainty limits, one may conclude that the first measurement was a statistically allowable event. However, if the excursion is repeated, appropriate investigative actions should be considered.

Most of the significant sources of uncertainty in radiochemical data are known to a laboratory and can be estimated. These include uncertainties associated with sample and background counting, radiochemical yield determination, efficiency calibration, and blank assessment. Other less easily defined but significant sources of uncertainty include those associated with self-absorption and quench correction, sample density correction, sample geometry variation, gamma photopeak area determination, determination of sample volume or weight, and dead time correction.

The uncertainty of a measured value is controllable, within certain limits, by decreasing the uncertainty associated with some input parameters. For samples containing low levels of radioactivity, a large component of the combined standard uncertainty may be associated with the instrumental assessment (counting) of the sample aliquant, i.e., the standard uncertainty of the net count (gross sample count minus background count). Increasing the total net count accumulated, or decreasing the uncertainty of the instrument background, or both, will decrease the counting uncertainty. Changes that may be made to decrease the counting uncertainty include increasing

the counting time for the sample or background, increasing the sample aliquant size (unless the sample geometry, quench, or self-absorption factors offset the gain in total radioactivity counted), using a more efficient geometry or detector, using an instrument with a lower background, and reanalyzing the sample to obtain a greater radiochemical yield. It also may be possible to concentrate the sample, which has the equivalent effect of increasing the sample aliquant size.

18.4 Radiochemistry Performance Indicators

Section 18.3 discussed how to evaluate radiochemistry performance indicators using statistically based control chart techniques. Any of the indicators below (blanks, replicates, laboratory control samples, matrix spikes, certified reference material, or tracer yield) can be evaluated using the control chart techniques. Analysts can observe individual *Z* score values to identify loss of control. Control charts will assist laboratory personnel in identifying the quality trends and excursions of any performance indicator.

18.4.1 Method and Reagent Blank

Issue: A method blank is a sample of a matrix as similar as practical to the associated samples that is free from the analytes (radionuclides) of interest to the extent possible. The method blank is processed simultaneously with, and under the same conditions as, samples through all steps of the analytical procedures. A reagent blank consists of the analytical reagent(s) in the procedure without the target analyte or sample matrix, introduced into the analytical procedure at the appropriate point and carried through all subsequent steps to determine the contribution of the reagents and of the involved analytical steps.

Blank samples are used to determine whether any radionuclide contamination is introduced by the measurement process. They assist in the control of any contamination introduced by the laboratory. Ideally, no target analytes should be present in the blank at detectable concentrations. If that is not possible (e.g., for naturally occurring radionuclides), those radionuclides should be extremely well-characterized and tracked. Control charts can be used to track these radionuclide levels in blanks. Using *X* charts, the laboratory can establish a program that evaluates the levels and trends of radionuclides in the different laboratory blanks. The techniques for establishing such a control chart program are described in Attachment 18A.

Discussion: The method blank is assumed to be representative of all samples in the batch with respect to the matrix and contamination assessment. When practical, it consists of the same or equivalent medium as the analytical samples, such as a deionized water blank for aqueous samples. Soil blanks are often prepared using “clean sand,” commercially available fine-grained

or beach sand whose inherent concentrations of target radionuclides are small and have been characterized sufficiently by the laboratory to allow its use as a blank. This approach may not be appropriate for very low-level analyses. Powdered, natural-matrix Standard Reference Materials (SRMs) are commercially available from National Institute of Standards and Technology (NIST) and also may be suitable (Section 18.4.5). However, due to the natural variability of soils, each choice of method blank medium must be evaluated by the laboratory prior to use. The results of method blanks are not used to correct sample activities but only to monitor for contamination.

Reagent blanks are matrix-independent and assess any contamination only from the reagents and lab-ware. They are used to correct sample activities for the contribution of naturally occurring radionuclides in the reagents, and used like method blanks, to check for unexpected contamination. When reagent blank results are used to correct sample activities, it is important that the blank results be carefully monitored using control charts.

It is common practice for some laboratories to add the reagents into a volume of deionized water equal to the sample volume, while other laboratories simply add the required reagents to an empty container and process it as an analytical sample. In either case, it should be noted that the reagent blank is not monitoring the entire analytical process. The fundamental issue for each laboratory is to decide on the appropriate reagent blank necessary to obtain the needed information on the measurement system. Considerable variability exists among laboratories in the use and preparation of reagent blanks.

In general, the reagent blank's concentration of analyte is expected to be small compared to that of the sample. However, for some low-activity environmental samples this may not be the case, and the correction becomes increasingly important as the concentration of the analyte in the sample approaches background concentrations. In these cases, care should be taken to accurately quantify the levels of radionuclides in the reagent blanks.

It is important to minimize radionuclide concentrations in the blanks and bring these levels under control. This is usually achieved through careful selection of reagents, maintaining laboratory and counting areas free from contamination, and by segregating high and low activity samples. Thorough documentation of all blank values is essential to allow for the application of statistical tests to evaluate potentially anomalous values and delineate their extent.

Ideally, the analyte concentration in a method or reagent blank should be as close to zero as possible, and replicate measurement of the blanks should be consistent within counting statistics. Acceptance criteria for blank results should be established and applied to all data, and should include warning and control limits (Section 18.3.2, "Statistical Means of Evaluating Performance

Indicators — Control Charts"). Blank values require scrutiny as part of the data evaluation and validation process for each analytical batch. Should restocking of reagents or other wholesale laboratory changes occur during a project, the method and reagent blanks prepared under the new conditions should be re-evaluated to ensure that they continue to be within established criteria.

An example of a numerical performance indicator for a method blank or a reagent blank used to monitor for unexpected contamination is

$$Z_{\text{Blank}} = \frac{x}{u_c(x)} \tag{1}$$

where x denotes the measured blank activity and $u_c(x)$ denotes its combined standard uncertainty. Warning limits for Z_{Blank} are ±2 and control limits are ±3. As mentioned earlier, if a reagent blank is used to blank-correct sample results, the blank results should be evaluated using control charts.

Typically, one method blank and/or reagent blank is analyzed with each batch or grouping of analytical samples regardless of batch size. Situations may occur where more frequent blanks are required to ensure that analytical conditions are stable, particularly when analyzing high and low concentration samples in the same analytical batch, or when instruments, reagents, or analytical method are suspect.

In general, corrective actions include procurement control of reagents, good laboratory cleaning practices, sample segregation according to anticipated concentrations, and instrument-related concerns, as discussed in this section. Good laboratory cleaning protocols should incorporate the evaluation of method and reagent blank performance to indicate if current practices are adequate. Instrument background data indicate a system's stability, and can be used to pinpoint the source of contamination, as can routine contamination (removable and fixed) surveys of laboratory and counting areas that are performed by the organization's health physics or radiation safety personnel.

Excursion: Blank changes can be grouped into three general categories: rapid changes, gradual increase or decrease, and highly variable changes. These are represented in Figure 18.2 and described below.

Rapid Changes: A sudden change in a blank value indicates the existence of a condition requiring immediate attention. Sudden changes often are caused by the introduction of a contaminant from high concentration samples, impure reagents, or contaminated sample preparation areas. Laboratory cleaning practices and new or recently restocked reagents

should be checked. When a sudden, significant increase in the blank occurs in conjunction with the introduction of new reagents through restocking or other changes, the causes should be investigated and if the reagent is contaminated, the reagent contributing the activity should be discarded and replaced. Particular attention should be paid to the samples counted directly prior to the contaminated blank, since small amounts of residues from these samples can contaminate the detector and have large effects on subsequent results when analyzing samples at or near environmental background. It may be necessary to take swipe or smear samples of questionable areas to identify the contaminant's source followed by a thorough cleaning or decontamination of all affected areas. Additionally, method or reagent blank values that are suddenly depressed should be investigated and may indicate other problems, including instrument malfunction like a loss of counting gas, incomplete chemical separation during the chemical preparation, or the failure to add necessary reagents. These other problems may be reflected in other areas, such as instrument performance checks or tracer yields.

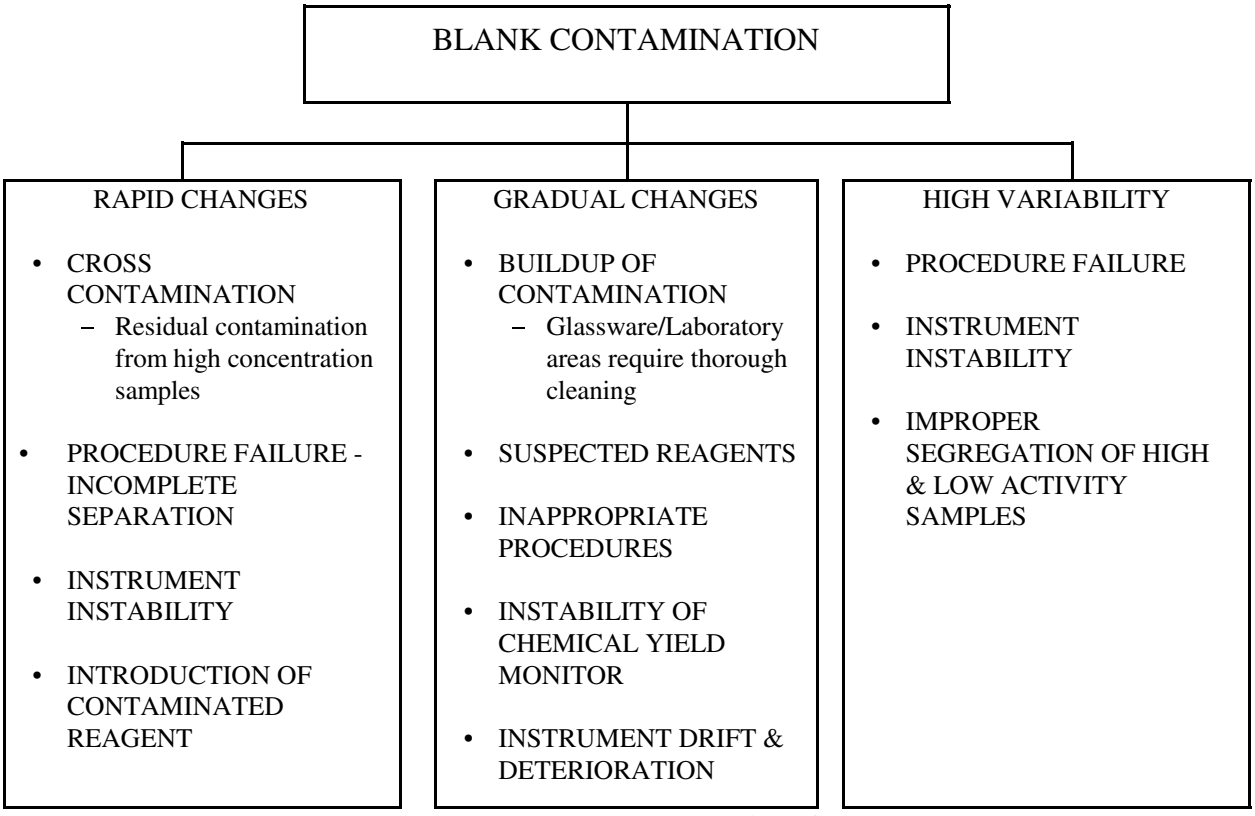

FIGURE 18.2 — Three general categories of blank changes

Gradual Changes: Gradually increasing blank values indicate the need to inspect all sample preparation and counting areas for sources of residual contamination. Often housekeeping or routine contamination control details such as cleaning glassware or instrument counting chambers are sufficient to bring blank values under control. Alternatively, gradually decreasing blank values warrant scrutiny with respect to proper instrument settings and procedural related problems like a lack of tracer/sample exchange, failure of chemical separation reactions, or the addition of all necessary reagents. The importance of documenting method and reagent blank values in this regard cannot be overemphasized, since data evaluation and trending analyses are impossible without complete records.

High Variability: Because method blank values are expected to be near zero, the degree of variability they exhibit should reflect the statistical variation inherent in radiometric determinations near these levels. Large variations in blank values typically indicate problems related to instruments or sample processing, as discussed in the two previous sections.

18.4.2 Laboratory Replicates

Issue: A laboratory replicate is two or more aliquants taken at the first subsampling event, normally after homogenization. In the event that there is no subsampling (when the method calls for using the entire sample) replicate analysis typically involves counting the prepared sample twice. The results of laboratory replicates are used to evaluate the precision of the measurement process. Note that counting a sample twice only assesses the instrument portion of the measurement process.

Precision is a measure of agreement among replicate measurements of the same property under prescribed similar conditions. Precision is a fundamental aspect of the analytical process and should be evaluated routinely as part of the laboratory's quality system. Evaluation typically is performed using multiple analysis of the same sample (blanks, spikes, blinds, reference materials, performance evaluation samples, etc.), in whole or part, and evaluating the analyses relative to a statistically based criterion. The range of sample types requires that the sample matrix's effects on the precision be captured and evaluated by the laboratory's routine quality control practices. The reproducibility of analytical results should be evaluated by replicates to establish this uncertainty component.

Discussion: The purpose for measuring precision is to determine whether the laboratory can execute an analytical method consistently and obtain results of acceptable variability. Analytical samples cover a range of physical forms or matrices, from homogeneous samples like finished drinking water to complex soils or heterogeneous wastes, and each matrix has the potential to

affect a protocol's precision.

In general, precision for aqueous samples tends to be less affected by sample heterogeneity than other media because if the sample's constituents are dissolved the sample is essentially homogeneous. This facilitates dividing the samples into equivalents fractions or aliquants. Multi-phase and high-solid-content samples that are heterogeneous are more problematic.

The acceptance criterion for precision should be related to the combined standard uncertainties of the measured results. The uncertainty of a result may depend on many factors (e.g., dissolved solids in water or particle sizes of soil), but such factors should affect the acceptance criterion only through their effect on the standard uncertainty.

As an alternative to sample duplicates, a matrix spike duplicate is sometimes used as an indicator of the analytical precision, as discussed in Section 18.4.3. A matrix spike duplicate is treated in the same manner as an unspiked replicate: both samples (original and duplicate) are processed identically to the other samples in the batch, and each aliquant is treated as an individual sample.

If the sample has multiple phases, the phases should be separated for individual analysis. For heterogenous materials, multiple analyses should be used, or the combined standard uncertainty of the results should be increased, to account for subsampling error (Appendix F). A typical frequency for replicate analyses is a minimum of one per analytical batch, regardless of batch size. Batch is defined as samples of similar matrix type with associated QC samples analyzed under the sample conditions at approximately the same time.

All analytical batches should be evaluated with respect to precision, whether by using replicates or matrix spike duplicates. This is done typically by the use of an acceptance criterion that derives a statistic that quantifies the difference between two values obtained by analyzing the same sample. Limits are then placed on the criterion, and data for any batch in excess of the criterion require investigation and corrective action as appropriate. An example of a numerical performance indicator for laboratory replicates is

$$Z_{\text{Rep}} = \frac{x_1 - x_2}{\sqrt{u_c^2(x_1) + u_c^2(x_2)}} \qquad (2)$$

where x_1 and x_2 denote the two measured activity concentrations and $u_c(x_1)$ and $u_c(x_2)$ denote their respective combined standard uncertainties. Warning limits for Z_{Rep} are ±2 and control limits are ±3.

Excursions: A regularly scheduled evaluation of precision with respect to the acceptance criterion should be an integral part of the laboratory quality system. Careful attention should be paid to the nature and anticipated analyte concentrations of all samples processed by the laboratory. Prospective identification of samples where precision is expected to be problematic often can address difficulties in this area. The choice of appropriate analytical method and analyst training are also important. An analyst needs to be familiar with specific steps in the procedure that provide an indication of incomplete processing.

Precision exhibits a range of values and depends in part on sample matrix and activity, assuming correct execution of the analytical method. Small changes, positive and negative, are expected and should be captured in the acceptance criterion's range. It is also sensitive to sample heterogeneity or errors in processing, such as incomplete chemical separation or sample dissolution, and lack of tracer or carrier equilibration. When performance indicators for precision are outside acceptance criteria, the laboratory should determine the reasons why and implement corrective actions.

Certain samples will exhibit higher variability because of their matrix, or the proximity of their analyte concentration to ambient background, as discussed previously. Consideration should be given to cases where a matrix requires the development and implementation of a specific acceptance criterion. The main causes for lack of precision (Figure 18.3) can be grouped as follows:

- Laboratory subsampling — subsampling techniques produced two dissimilar aliquants from one sample, and the original and duplicate are not the same. An analyst should be careful to ensure that the sample is thoroughly homogenized before subsampling.

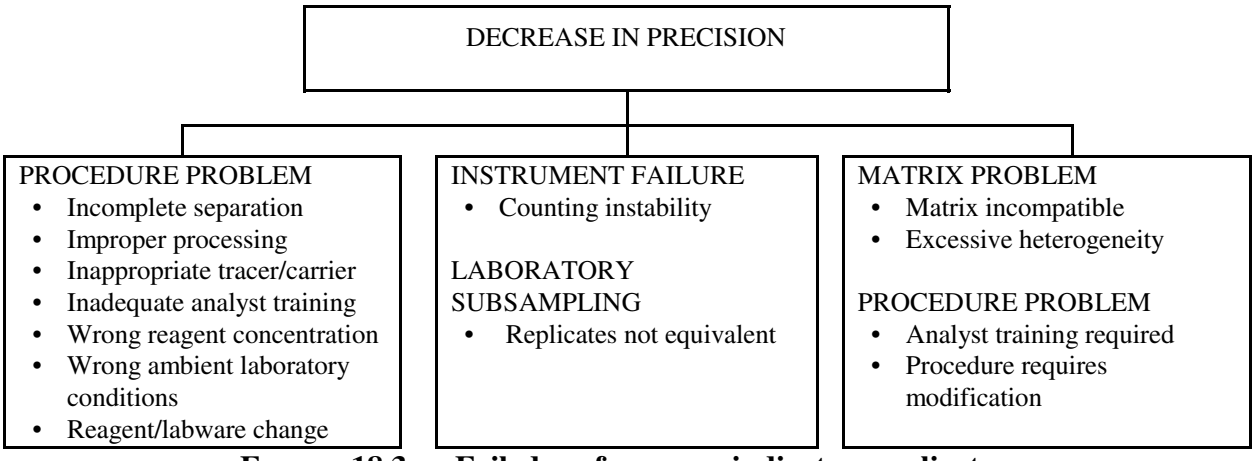

FIGURE 18.3 — Failed performance indicator: replicates.

- Matrix – Sample constituents interfere with preparation chemistry, e.g., coprecipitation of interfering non-analyte radionuclides from sample or excessive dissolved solids.

- Counting statistics – Sample activity is so low that small statistical variations in background cause disproportionate responses.

- Contamination – Intermittent contamination from measurements system, glassware, etc., produces anomalous data for the original sample, but not the duplicate/replicate.

- Other – Failed chemical process, failed instrumentation, training, failed lab environment, failed procurement control.

18.4.3 Laboratory Control Samples, Matrix Spikes, and Matrix Spike Duplicates

Issue: A laboratory control sample (LCS) is a QC sample of known composition (reference material) or an artificial sample, created by fortifying a clean material similar in nature to the environmental sample. The LCS is prepared and analyzed in the same manner as the environmental sample. A matrix spike (MS) is an aliquant of a sample prepared by adding a known quantity of target analytes to a specified amount of sample and subjected to the entire analytical procedure to establish if the method or procedure is appropriate for the analysis of the particular matrix. A matrix spike duplicate (MSD) is a second replicate matrix spike prepared in the laboratory and analyzed to evaluate the precision of the measurement process.

An important performance indicator is the ability to ensure that the analytical methods employed obtain data that are representative of the true activity in a sample, i.e., produce data that are accurate. The routine analysis of spiked samples provide data for an evaluation of the laboratory's reported measurement uncertainty and allow for the determination of bias, if one exists. Evaluation is typically performed using prepared samples consisting of media equivalent to a routine analytical sample with a known, measurable amount of the analyte of interest. Upon completion of the analysis, the results are compared to the known or accepted value, and the agreement is evaluated using a predetermined criterion. The range of sample types assayed in a laboratory may require that spikes are prepared using several sample media. Use of matrix spiked samples will reflect the analytical method's ability to make accurate quantitative determinations in the presence of the matrix.

Discussion: As stated previously, analytical samples cover a range of physical forms or matrices, and each matrix can change a method's expected bias. Tracking sets of LCS and matrix spike results can give laboratory personnel an indication of the magnitude of bias. Care must be taken

when analyzing site specific matrix spike results because these matrices may be very complex and subject to large variability. In general, aqueous samples tends to be less affected than other media like soils or heterogeneous materials. However, multi-phase fluids, high solid content, and brackish or saline waters may be more problematic.

The analyst should carefully consider the spiking levels for laboratory control samples and matrix spikes. Spikes and LCSs may be prepared near the lower limits of detection to test the methods performance on clean or slightly contaminated samples. Conversely, matrix spikes and LCSs may be spiked at high levels for groups of highly contaminated samples. The laboratory should try to spike at or near the action level or level of interest for the project.

Possible numerical performance indicators for laboratory control samples and matrix spikes are

$$Z_{\mathrm{LCS}} = \frac{x - d}{\sqrt{u_c^2(x) + u_c^2(d)}} \tag{3}$$

$$Z_{\mathrm{MS}} = \frac{x - x_0 - d}{\sqrt{u_c^2(x) + u_c^2(x_0) + u_c^2(d)}} \tag{4}$$

where x is the measured value of the spiked sample, d is the spike concentration added, x_0 is the measured concentration of the unspiked sample, and $u_c^2(x)$, $u_c^2(d)$, and $u_c^2(x_0)$ are the squares of the respective standard uncertainties. The warning limits for either of these indicators are ±2 and the control limits are ±3.

Excursions: Excursions in the LCSs and MSs can be used to identify various out of control situations. The advantage to the LCS is that the sample matrix is always the same so matrix effects should not be a factor in evaluating excursions. A rapid and one-time excursion in the LCS usually indicates that a mistake was made in the procedure. A rapid change with continued occurrences suggest that something occurred that is out of the ordinary, such as a new analyst performing the procedure or a new standard solution or new reagents being used. If an LCS shows elevated concentrations, analysts should check for contamination sources or poorly prepared spiking solutions. Slow changes showing a trend usually indicate degradation or contamination of equipment or reagents and may be indicative of bias and should be investigated.

Excursions of MSs can be difficult to interpret if the matrix changes from batch to batch. However, an excursion may indicate that the method is not appropriate for a particular matrix. If

the MS shows lower than expected concentrations, the analyst should check for poor techniques or expired or poorly prepared reagents and spiking solutions.

Elevated or depressed results for site-specific MSs need to be interpreted with the results from LCSs. If both the LCS and site-specific MS results are elevated or depressed then the cause is usually internal to the laboratory. If only the site-specific MS is depressed or elevated, the cause usually is due to the matrix.

18.4.4 Certified Reference Materials

Issue: Certified reference materials (CRMs) are well-characterized, stable, homogeneous materials with physical or chemical properties determined within specified uncertainty limits. Laboratories that analyze CRMs can compare their performance to the certified concentration and uncertainty levels. CRMs are used for the calibration of an apparatus or the assessment of a measurement method.

Discussion: Metrology organizations issue CRMs in various matrices with critically evaluated concentration values for the radionuclide constituents. A CRM issued by NIST or under license from NIST is called a "standard reference material" (SRM). The usefulness of a reference material depends on the characterization of the radionuclide source, activity levels, and their estimated uncertainties.

CRMs can be used as internal laboratory QC samples to evaluate the ability of analytical methods to handle the matrix. CRMs need not be known to the analyst but can be introduced into the analytical stream as a blind. Comparison of analytical results of CRMs to their certified values provides linkage to the national scale of measurements and a measure of method accuracy.

The planning that goes into the preparation of a CRM involves the selection of analytical techniques that have adequate sensitivity and precision for specific analyses. It has become increasingly important to have available well-characterized CRMs of a natural "matrix" type, which may be used in laboratory tests of measurements of environmental radioactivity. Such materials may be used in the evaluation of competing analytical methods, and also in the cross-comparison of interlaboratory data—both at the national level and the international level.

The Ionizing Radiation Division of NIST has constructed several SRMs for radiation measurements. These are included in the 4350 series and can be ordered through NIST. One widely used SRM is the natural matrix ocean sediment (4357). The radionuclides in the NIST natural matrix SRMs are not spiked into the matrix but are incorporated through natural

processes to present the analyst with the combination of species that may be faced on a routine basis. The SRM 4357 has two sediment sources: the Chesapeake Bay (benign) and the Irish Sea ("hot").

The NIST natural matrix SRM project has certified actinides, fission and activation radionuclides in soils, freshwater lake and river sediments, human tissues, and ocean sediment, and is working on additional unique matrices: ashed bone, ocean shellfish, and Rocky Flats Soil-II.

A numerical performance indicator for the analysis of a CRM is essentially the same as that for a laboratory control sample. An example is

$$Z_{\mathrm{CRM}} = \frac{x - d}{\sqrt{u_c^2(x) + u_c^2(d)}} \tag{5}$$

where x is the measured value, d is the certified value, and $u_c^2(x)$ and $u_c^2(d)$ are the squares of the respective combined standard uncertainties. Warning limits for Z_{CRM} are ±2 and control limits are ±3.

Excursions: Excursions in the CRM results can be used to identify various out-of-control situations. The advantage of the CRM is that the sample matrix is always the same, and the levels of analytes are known to a high degree, so uncertainties in matrix effects and radionuclide content should not be a factor in evaluating excursions. A rapid and one-time excursion in the SRM usually indicates that a mistake was made in the procedure. A rapid change with continued occurrences suggest that something occurred that is out of the ordinary, such as a new analyst performing the procedure or the use of a new batch of calibration solutions or reagents. Slow changes showing a trend usually indicate degradation or contamination of equipment or reagents.

If a CRM result shows elevated concentrations, analysts should check for contamination sources or poor instrument calibration. If the results show decreased concentrations, the analyst should check for poor techniques or expired or poorly prepared reagents and solutions.

CRM results may indicate a bias in the measurement process. Tracking the performance of several consecutive CRM measurements will show if the method or the laboratory consistently obtains high or low results. If the results are consistently higher or lower than the certified values, they should be evaluated for a statistical difference, e.g., *t*-tested. When the test indicates a statistical difference, a bias is indicated and the laboratory should investigate the cause of the bias and correct or characterize it.

Example: The NIST ocean sediment SRM 4357 offers a good example of a material for

evaluating a laboratory performance using a specific analytical method. The blended sediment sample has been analyzed by a number of laboratories, and 10 radionuclides have certified activity values (Lin et al., 2001). The six "natural" radionuclides concentrations tended to have normal distributions (Table 18.2a), while the four "man-made" radionuclides tended to have Weibull distributions (Table 18.2b). There are also 11 other radionuclides where the activity concentrations are not certified at this time but may be at some future time (Table 18.2c).

TABLE 18.2a — Certified Massic activities for natural radionuclides with a normal distribution of measurement results

Radionuclide	Mean ± $2s_m$ ($mBqg^{-1}$)	Tolerance Limit (2.5 to 97.5%) ($mBqg^{-1}$)	Number of Assays	Half-Life ± 1s (In years)
^{40}K	225 ± 5	190 – 259	31	$(1.277 \pm 0.008) \times 10^9$
^{226}Ra	12.7 ± 0.4	10.3 – 15.0	21	1600 ± 7
^{228}Ra	13.3 ± 0.8	9.2 – 17.4	20	5.75 ± 0.03
^{228}Th	12.1 ± 0.3	9.7 – 14.6	40	1.9131 ± 0.0009
^{230}Th	12.0 ± 0.5	9.6 – 14.4	18	75380 ± 300
^{232}Th	13.0 ± 0.3	11.6 – 14.3	18	$(1.405 \pm 0.006) \times 10^{10}$

Table 18.2b — Certified Massic activities for anthropogenic radionuclides with a Weibull distribution of measurement results

Radionuclide	Mean ± $2s_m$ ($mBqg^{-1}$)	Tolerance Limit (2.5 to 97.5%) ($mBqg^{-1}$)	Number of Assays	Half-Life ± 1s (In years)
^{90}Sr	4.4 ± 0.3	2.1 – 8.4	49	28.87 ± 0.04
^{137}Cs	12.7 ± 0.2	10.8 – 15.9	76	30.07 ± 0.03
^{238}Pu	2.29 ± 0.05	1.96 – 2.98	65	87.7 ± 0.3
^{239}Pu + ^{240}Pu	10.4 ± 0.2	9.3 – 13.2	84	24110 ± 30 6564 ± 11

Table 18.2c — Uncertified Massic activities. Radionuclides for which there are insufficient data or for which discrepant data sets were obtained. Uncertainties are not provided because no meaningful estimates could be made.

Radionuclide	Mean ($mBq\ g^{-1}$)	Range of Reported Results ($mBq\ g^{-1}$)	Number of Assays	Half-Life ± 1s (In years unless listed as minutes, hours, or days)
^{129}I	0.009	0.006 – 0.012	6	$(1.57 \pm 0.04) \times 10^7$
^{155}Eu	1.4	1.2 – 1.5	2	4.68 ± 0.05
^{210}Po	14	12 – 15	5	138.376 ± 0.002 d
^{210}Pb	24	14 – 35	19	22.3 ± 0.2
^{212}Pb	14	13 – 14	5	10.64 ± 0.01 h
^{214}Bi	15	9 – 20	5	19.9 ± 0.4 m

Radionuclide	Mean (mBq g^{-1})	Range of Reported Results (mBq g^{-1})	Number of Assays	Half-Life ± 1s (In years unless listed as minutes, hours, or days)
^{234}U	12	9 – 15	68	$(2.45 \pm 0.02) \times 10^5$
^{235}U	0.6	0.1 – 1.4	63	$(7.038 \pm 0.006) \times 10^8$
^{237}Np	0.007	0.004 – 0.009	9	$(2.14 \pm 0.01) \times 10^6$
^{238}U	12	7 – 16	76	$(4.468 \pm 0.003) \times 10^9$
^{241}Am	10	7 – 18	97	432.7 ± 0.6

SRM 4357. Data for these radionuclides are provided for information only. The Massic activities are not certified at this time, but may be certified in the future if additional data become available.

18.4.5 Chemical/Tracer Yield

Issue: Some methods require that radionuclides should be separated chemically from their sample matrix and purified before measurement. During chemical processing, some of the analyte radionuclide will be lost due to sample spillage, evaporation, incomplete chemical reactions (i.e., precipitation or extraction), etc., as discussed in Chapter 12. While these losses may correlate with a group of samples of similar chemical composition or from the same sampling area, they can be sample specific. For quantitative analysis, it is necessary to correct observed instrument responses for these losses for each analytical sample. Corrections are made using compounds that are stable (carriers) or radioactive (tracers). An inappropriate method for determining chemical yield may result in an analytical bias.

Discussion: Most alpha- and beta-emitting radionuclides require chemical separation prior to measurement, in part because of the short effective range of the radiation.

CARRIERS. Since it is impossible to determine exactly how much of the analyte is lost during processing, and because the physical mass of the radionuclide is too small to measure gravimetrically, a compound is added to the sample at the start of the chemical processing, and is carried through the analytical process and assayed. The added compound typically is stable and exhibits the same chemical properties as the analyte and therefore "carries" the analyte radionuclide—for example, stable barium that carries radium isotopes, or stable yttrium that carries ^{90}Y. These added compounds are called "carriers" and are added in sufficient quantity to allow gravimetric assay upon completion of the analysis. The ratio of the carrier recovered to the amount added is the chemical recovery, or yield. Because the carrier and analyte exhibit similar chemical behavior, the chemical yield of both should be equal, i.e., if 85 percent of the stable barium is recovered, then it follows that the observed instrument response represents 85 percent of the radium present in the sample.

TRACERS. For radionuclides above atomic number 83, stable isotopes do not exist, and a different approach is taken to determine the analyte's yield. For these radionuclides, an isotope other that those being measured is added to the sample in the same manner as described above, e.g., ^{232}U used as a tracer for isotopic uranium (^{234}U, ^{235}U, and ^{238}U), ^{236}U, or ^{242}Pu used as a tracer for isotopic plutonium (^{238}Pu, ^{239}Pu, and ^{240}Pu).

This approach to chemical yield determination is based on the following assumptions regarding the carrier/tracer:

- It exhibits similar chemical behavior as the analyte under the protocol's conditions.
- The energy emission of the tracer and progeny should not interfere with the resolution of the analytes of interest.
- It is chemically and physically equilibrated with the sample before losses of either occur.
- Indigenous concentrations of carrier or tracer are insignificant, or are well known and can be quantified and corrected for during subsequent data analysis.
- The chemical form of carrier or tracer precipitates are consistent with what was used during the material's preparation and standardization.

Care should be taken during the analytical procedure to ensure that these assumptions are valid. Different conditions, such as a lack of equilibrium between the tracer and sample analyte, can result in inaccurate data. If there is indigenous tracer or carrier in the sample, this quantity should be known so that the appropriate correction can be made for its contribution to the chemical yield. In some cases, this will prevent the procedure's use, as described below. As stated previously, the quantity of tracer or carrier added to the sample should overwhelm its indigenous concentration, which cannot be determined for samples with unknown tracer or carrier content. A separate analysis for trace elements or interfering radionuclides could provide information to estimate the uncertainty contributed by the sample's indigenous tracer or carrier.

It should be noted that some analytical methods exclude direct assessment of the procedure's chemical recovery for each sample analysis, e.g., *Procedure 908.1 for Total Uranium in Drinking Water* (EPA, 1980b). In such cases, chemical recovery is typically addressed by analyzing a group of prepared standards by the same protocol and the results are analyzed statistically to derive a chemical recovery factor. The recovery factor is applied to routine samples based on the assumption that the standards used for its derivation are representative of routine samples. This

approach precludes the empirical assessment of a sample specific chemical recovery, and would probably require scrutiny and periodic verification.

Acceptance limits for chemical/tracer yields should be specified in the laboratory's Quality Manual. While it is customary to establish lower limits for chemical yield, upper limits may also be necessary since excessive yields indicate a loss of analytical control. All limits developed by the laboratory should be either statistically based or based on historical data, and should include warning and control limits. The inherent differences among sample matrices generally require the use of matrix specific criteria, i.e., finished drinking water limits may differ from limits for high solid content waters, sandy soils or heterogeneous media. Irrespective of medium, where practical, the chemical yield and its uncertainty should be determined, recorded and tracked for each radiochemical measurement.

Excursions: There are several possible reasons for the yield to be outside of the acceptance limits. These are summarized in Figure 18.4 and discussed below.

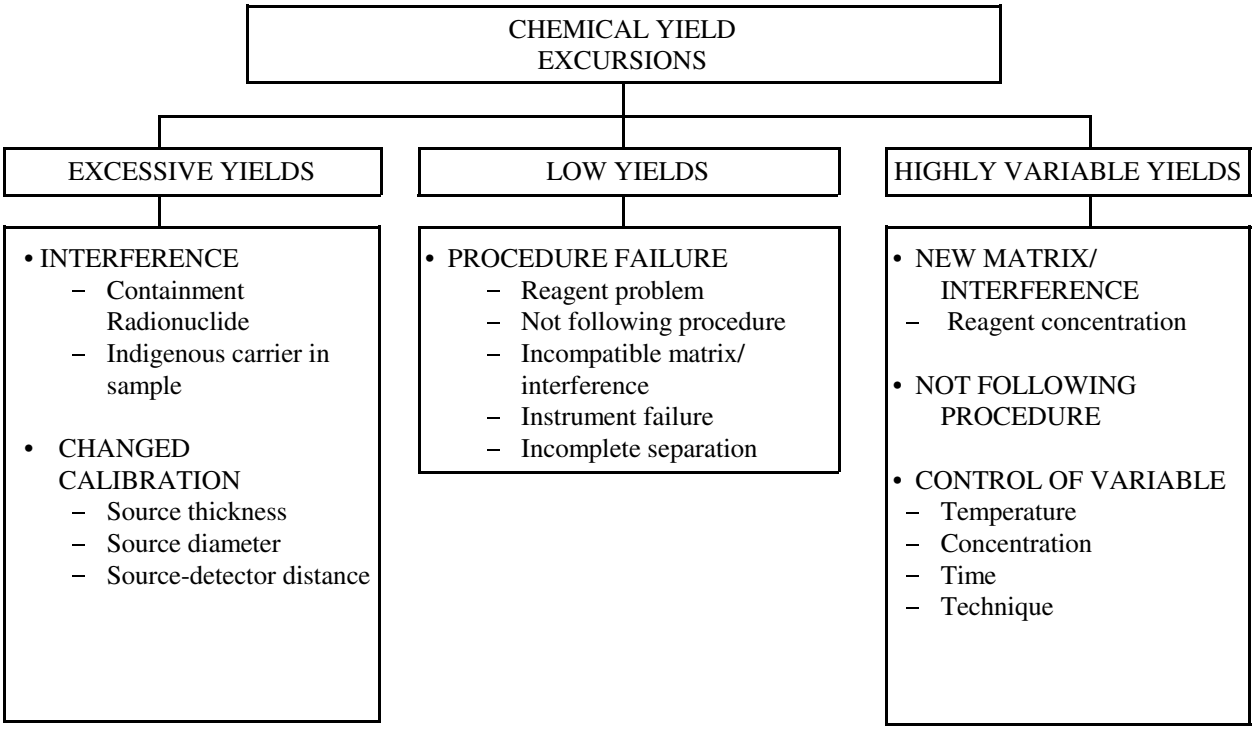

FIGURE 18.4 — Failed performance indicator: chemical yield

EXCESSIVE YIELDS: A chemical yield significantly greater than 100 percent indicates a problem. Typical causes of excessive chemical yields are provided below:

- Interference. The sample may contain an interfering radionuclide that cannot be distinguished from the tracer and therefore biases the tracer response; the sample may contain an indigenous concentration of the tracer or carrier used; or large amounts of another stable element are present.

- Counting. Changes in instrument calibration factor or other factors that affect counting, e.g., source thickness, diameter, source-detector distance or change in chemical form of final sample precipitate.

- Instrument failure.

LOW YIELDS: A very low yield usually indicates a procedural failure caused by incomplete or unsuccessful chemical separation, matrix interference, missing reagents, or the exclusion of a key element in the sample processing. A significantly lower yield will increase the overall measurement uncertainty and degrade the procedure's effective detection capability unless the counting time is appropriately extended, which may be impractical or even ineffective in many cases. Furthermore, measurement of the recovered carrier or tracer becomes increasingly more adversely affected by background, stable element, water absorption, and other corrections as the yield decreases. Fixed lower limits for yields often are established and should be specific to analytical procedures and sample matrices. Setting an upper limit is recommended for the acceptable relative uncertainty in a yield measurement.

HIGHLY VARIABLE YIELDS: High variability in procedural temperature, concentration, time, reagent concentration, or laboratory technique can have dramatic effects on yield. Highly variable yields indicate a lack of procedural control and should be investigated and corrected. A simple step such as heating samples on a hotplate can lead to variability in yield because the hotplate surface is thermally uneven. Samples can be dried and reconstituted several times during the course of the preparation protocol, and samples may require different amounts of heat or water, which introduces additional variability. When highly variable chemical yields are observed, a careful examination of the analytical procedure's application is recommended to determine critical variables and the controls needed to re-establish adequate management over yields.

18.5 Instrumentation Performance Indicators

Radiometric and non-radiometric instruments are used currently to quantify radionuclides in a variety of environmental matrices, and quality control measures are necessary to ensure proper instrument performance. This section presents radiometric instrument performance measures that indicate a measurement system is in control. For detailed information on instrument concepts and specific techniques, see Chapters 15 and 16 as well as ASTM standard practices (e.g., D3648, for the Measurement of Radioactivity). The specific quality control procedures to be followed depend on the measurement equipment. Sufficient checks are needed to demonstrate that the measurement equipment is properly calibrated, the appropriate background has been recorded, and that all system components are functioning properly. QC measures for instrumentation should include at a minimum: (1) instrument background measurements, (2) instrument calibration with reference standards, and (3) periodic instrument performance checks subsequent to the calibration. Acceptable control limits should be specified in the laboratory Quality Manual.

18.5.1 Instrument Background Measurements

Issue: In general, radionuclide detection covers more than 17 orders of magnitude of sample activity, from irradiated material that produces high radiation fields to environmental samples. All radiation detection instruments have a background response even in the absence of a sample or radionuclide source. To determine the instrument's response to the radioactivity contributed by the sample alone (net), the instrument background response is subtracted from the sample-plus-background response (gross). For discussions on possible contamination, refer to Section 18.4.1. Background corrections become more critical when the instrument net response is small relative to the background. Careful control of contamination and routine monitoring of instrument background are therefore integral parts of a control program. Inappropriate background correction results in analytical error and will increase the uncertainty of data interpretation.

Discussion: Every radionuclide detector produces a signal response in the absence of a sample or radionuclide source. These signals are produced by electronic dark current, cosmic radiation, impurities in the instrument construction materials, crosstalk between the detector's alpha and beta channels, sources in the general vicinity of the detector, and residual contamination from previous counting episodes. The majority of these contributors to instrument background produce a fairly constant count rate, given sufficient measurement time (i.e., dark current, cosmic radiation, construction material impurities). For other sources, instrument backgrounds vary as a function of time (i.e., from decay or ingrowth of residual contamination or as radon levels fluctuate throughout the day and season). For low-level measurements, it is imperative that the

background be maintained as low as feasible. Active or passive detector shielding, removing or adequately shielding radioactive sources in the vicinity of the detector, and good laboratory practices to prevent residual contamination are necessary to maintain low instrument background.

The instrument's background should be determined in the absence of a radionuclide source.The instrument background should be well characterized. The instrument background is an important factor in determining the ability to achieve a specific minimum detectable concentration (MDC). Control limits for the background should be specified in the laboratory's Quality Manual, as appropriate. The background population considered in the statistical calculations should cover a sufficient period of time to detect gradual shifts in the measurement system's background contamination or detector instability. Additionally, backgrounds should be determined in such a way that they mimic actual sample measurement conditions as closely as possible, i.e., using appropriate sample containers, geometries, and counting times.

Background measurements should be made on a regular basis and monitored using control charts. For instruments with well established background performance records and a low probability of detector contamination, this frequency may be modified by the laboratory. For mass spectrometry and kinetic phosphorimetry analysis, background measurements should be performed on a real time basis. See ASTM E181, ANSI N42.12, and NELAC (2000) *Quality Systems Appendix D* for more information on the suggested frequency of background measurement.

Excursions: Variations in instrument backgrounds may indicate instrument malfunction. Variations may take the form of rapid increase or decrease in background, slow increase or decrease in backgrounds, and highly variable or erratic backgrounds. These variations can result in the measurement system's reduced precision and decreased detection capability. Rapid or significant increases in background measurements may be due to instrument or blank contamination, insufficient shielding with relocation of nearby radionuclide sources, or large scale equipment malfunction (e.g., a broken window on a gas proportional system).

Instrument background data should be evaluated for trends, which is facilitated by regular observation of control charts. A slowly changing background could alert laboratory personnel to a potentially serious instrument failure. A sufficient number of data points (Chapter 15) taken over time should be included in any trend analysis. Slowly changing instrument backgrounds could be caused by low counting-gas flow rates, small incremental instrument contamination, or electronic drift or noise.

When the instrument background is more variable than expected, the reliability of measurements becomes questionable, resulting in loss of confidence and increased uncertainty. This indicates a

loss of control over the measurement environment, or limitations of the data handling software. The root cause of the variability should be identified and corrected to re-establish statistical control over the instrument background. Table 18.3 presents reasons for changing backgrounds.

TABLE 18.3 — Instrument background evaluation

Instrument Background Failed Performance Indicator		
Rapid Change in Background	**Slow Change in Background**	**Excessively Variable Background**
Electronic failure	Instrument contamination	Sources being moved
Detector failure	Electronic drift	Radon fluctuation
Loss of coolant/vacuum	Low counting gas flow rate	Insufficient shielding
Instrument contamination		Insufficient counting statistics
Counting gas changes		Interfering radionuclides
Temperature/humidity fluctuation		Poor peak deconvolution
Laboratory contamination		Intermittent electrical short
External sources		Failing electronics
Insufficient shielding		
Personnel with nuclear medicine dose		

18.5.2 Efficiency Calibrations

Issue: This section discusses selected aspects of instrument calibration that are pertinent to laboratory quality control. A more in-depth, technical discussion is provided in Chapter 16. The number of events (counts) recorded by a detector is converted to activity (actual radionuclide transformations) by empirically determining this relationship with NIST-traceable radionuclide sources when available. This relationship is expressed in the system's efficiency calibration. A separate efficiency is determined for each detector-source combination and is typically energy or radionuclide specific.

Detector efficiency is critical for converting the detector's response to activity. As discussed above, routine performance checks can evaluate several aspects simultaneously (sample geometry, matrix, etc.) and provide a means to demonstrate that the system's operational parameters are within acceptable limits. These are typically included in the assessment of the analytical method's bias and are specified in terms of percent recovery based on the source's known disintegration rate. Performance checks for measurement efficiency are usually determined statistically based on repeated measurements with a specific check source. Detection of a shift in measurement efficiency should be investigated.

The frequency of performance checks for efficiency calibrations is instrument specific. The frequency of these checks is often based on a standardized time scale or a percentage of the total

number of analyses performed using that method.

Performance checks for instrument efficiency typically are performed on a day-of-use basis. The level of activity in the check source should be sufficient to allow the accumulation of enough counts in a short time so that daily performance checks do not impose an unnecessary burden on the laboratory. However, the source strength for spectrometry systems should be such that instrument dead time is not significant and gain shifts do not occur (ANSI 42.23). For detectors that are used infrequently, it may be necessary to perform a check before and after each set of measurements.

Control charts provide a useful tool for documenting and evaluating performance checks for efficiency calibrations, and should be established and maintained for the intrinsic efficiency of each detector. There are several methods available for evaluating performance using control charts (see Attachment 18A).

Discussion: Most radiation detectors do not record all of the nuclear transformations that occur in samples undergoing measurement, i.e., they are not one hundred percent efficient. This occurs for several reasons, and the prominent reasons are discussed briefly below.

- Intrinsic or absolute efficiency[2] – In the absence of all other factors, a detector will only record a fraction of the emissions to which it is exposed due to its composition and other material-related aspects. Intrinsic efficiency is a measure of the probability that a count will be recorded when a particle or photon of ionizing radiation is incident on a detector (ANSI N1.1).

- Geometry – The spatial arrangement of sample, shielding, and detection equipment, including the solid angle subtended by the detector and sample configuration, largely determines what fraction of the emissions from the source actually reach the detector (ANSI N15.37). Geometry includes the source's distance from the detector and its spatial distribution within the counting container relative to the detector and shielding components.

- Absorption – Radiation emitted by the sample can be absorbed by the sample itself (self

[2] Efficiency measures the fraction of emitted photons or particles that are actually detected. It is affected by the shape, size, and composition of the detector as well as by the sample-to-detector geometry. There are two ways that efficiency can be expressed: "Absolute efficiency" is the fraction of all the photons or particles emitted by the source that are actually detected, and "intrinsic efficiency" is the ratio of photons or particles detected to the number that actually fall on the detector.

absorption), as well as other materials placed between the source and the detector, i.e., sample container, detector housing and shielding (NCRP 58).

- Backscatter – Radiation emitted by the sample can hit the sample container and scatter into the detector.

The detector response is a composite of these factors.

Each radiation detector should be calibrated to determine the relationship between the observed count rate of the detector and the disintegration rate of the source being assayed. This relationship is called the efficiency calibration—typically expressed in counts per second/ disintegration per second, or cps/dps—and is an integral part of the measurement protocol. For alpha spectrometry systems, the efficiency of detection is energy-independent. Efficiencies for gamma spectrometry are energy dependent, and an efficiency calibration typically covers a range for a specific counting geometry, e.g., 50 to 1,800 kilo electron volts (keV).

Once this relationship is established, it should be checked at regular intervals using what is called a performance or calibration check. The performance check does not seek to reestablish the detector's efficiency but simply demonstrates that the relationship is within acceptance limits. When designed properly, an efficiency performance check evaluates the intrinsic efficiency, geometry and absorption in a single measurement. Accordingly, it takes the form of a single value that incorporates all effects for a target radionuclide and a specific detector-sample configuration. Detectors that are energy dependent and measure radionuclides with multiple energies, such as photon or alpha spectrometers, should have performance checks at several energies throughout the measurement range. For these detectors, the performance check can simultaneously address the system's efficiency, energy calibration and resolution using a single source. An internal pulser can be used to check the electronics.

Because the performance check's purpose is to demonstrate that the system's efficiency remains constant, the source's absolute disintegration rate need not be known, provided its purity can be established, its half-life is known, and its activity is sufficient to provide adequate precision. Accordingly, it is not necessary to use a NIST-traceable check source for this purpose. Check sources that are non-NIST-traceable can meet the precision objectives of the performance check and they are less expensive.

Excursions: Changes in the efficiency of a detector can only be corrected by determining the root cause of the problem and repeating the efficiency calibration. Gradual changes in geometry usually indicate a problem with the technique of sample mounting or preparation. A visual

inspection of the prepared sample is often helpful in eliminating sample geometry as a source of the problem. For example, a precipitated sample counted on a gas proportional counter has an expected appearance, i.e., a circle of precipitate centered on the planchet and often covered with thin plastic film. If the prepared sample does not have the correct appearance, there could be a problem with the geometry, self-absorption, and backscatter. This can sometimes be corrected by preparing the sample a second time, inspecting it and presenting it for counting a second time. Re-training personnel responsible for the error may also be indicated. Because samples that have been improperly prepared for counting can result in contamination of or physical damage to the detector, it is strongly recommended that every sample be visually inspected prior to counting. Significant changes in geometry caused by modifications to the source preparation method can only be corrected by recalibrating the detector. Examples of modifications to source preparation methods are (1) using a new filter so that the geometry of the test source is different than the geometry used for calibration, and (2) replacing the containers used for gamma spectrometry with containers that have a different wall thickness or are made from different materials.

Changes in intrinsic efficiency generally result from a physical change to the detector and often result in rapid changes in efficiency. In many cases, changes that affect the intrinsic efficiency of a detector render it inoperable. These are specific to a detector type and are listed below:

- HPGe, Ge(Li), and surface barrier detectors – Real or apparent changes in intrinsic efficiency caused by vacuum leaks or failure of field effect transistor.

- Thin window detectors (gas proportional counters, low-energy photon) – Changes in measurement efficiency are typically associated with damage to the detector window.

- Gas proportional systems – Problems with efficiency related to the quality or flow of counting gas.

- Anti-coincidence systems with guard detectors – Electrical problems with the anti-coincidence circuits that may produce apparent changes in efficiency.

- Scintillation detectors – Gradual changes in efficiency are associated with the scintillator or the photomultiplier tube. For example, NaI(Tl) crystals may gradually turn yellow over time resulting in a lower intrinsic efficiency, and liquid scintillation counters may have residue gradually build up on the surface of the photomultiplier tube affecting the detection of photons by the tube.

18.5.3 Spectrometry Systems

18.5.3.1 Energy Calibrations

Issue: This section discusses selected aspects of instrument calibration that are pertinent to laboratory quality control. A more in depth, technical discussion is provided in Chapter 16. All radiation measurements are energy dependent to a certain extent. However, spectrometric techniques such as gamma and alpha spectrometry identify radionuclides based on the energy of the detected radiations. For these techniques a correct energy calibration is critical to accurately identify radionuclides. Problems with energy calibration may result in misidentification of peaks.

Discussion: Spectrometry systems should be calibrated so that each channel number is correlated with a specific energy. To identify radionuclides correctly, this energy calibration needs to be established initially and verified at regular intervals. The energy calibration is established by determining the channel number of the centroid of several peaks of known energy over the applicable energy range. Typically, a minimum of three peaks is used, and commercially available sources contain nine or ten photopeaks. The relationship between energy and channel number can be determined by a least squares fit. To account for non-linearity, a second or third order fit may be used. However, these require more points to define the curve. For example, a first order calibration requires at least two points, while a second order calibration requires a minimum of three points. The end points of the curve define a range of applicability over which the calibration is valid, and peaks identified outside the curve's range should be used carefully. The uncertainty associated with the curve should be available at any point along the calibration curve.

Quality control checks for energy calibration may be combined with checks for efficiency calibration and resolution. Radiations emitted over the range of energy of interest are measured, and two or more peaks are used to demonstrate that the energy calibration falls within acceptable limits. Check sources may consist of a single radionuclide (e.g., ^{137}Cs or ^{60}Co) or a mixture of radionuclides (e.g., mixed gamma). Because only the location of the peak is of concern, there is no requirement that the check source be calibrated or certified, except for ensuring that it does contain the radionuclide(s) of interest at a specified level of purity.

The energy calibration is determined when the system is initially set up by adjusting the gain of the amplifier, analog-to-digital conversion (ADC) gain, and zero. Criteria that indicate when readjustment is required because of gradual and abrupt changes in the energy versus channel calibration should be established as an integral part of the system's operating procedure. These changes usually are monitored by the measurement system's software, and the user specifies the

allowable difference between that the system's response and the radionuclide's known energy. The tolerable difference often relates to the instrument's resolution. For example, a high resolution instrument such as an intrinsic germanium detector typically will have acceptable limits on the order of a few keV, while a low resolution instrument such as a NaI(Tl) detector typically will have acceptable limits on the order of several tens of keV.

Spectra also can be analyzed by identifying each peak manually. With manual identification, the acceptable limits for the energy calibration are determined for each spectrum based on the professional judgment of the person analyzing the spectrum.

The frequency of QC checks for energy calibrations can be related to the expected resolution of the instrument, the electronic stability of the equipment, or the frequency needs of QC measurements for efficiency calibration or resolution. These are specified typically in the laboratory's Quality Manual or other typical project-related documentation. Examples for three detector types are provided below and in Table 18.5.

- HPGe and Ge(Li) Photon Detectors. Energy calibrations are typically verified using a check source on a day of use basis. Every sample spectrum should include verification of the energy calibration as part of the data review process, when possible. Under extreme conditions (e.g., in situ measurements in bad weather), it may be necessary to perform checks at the beginning and end of each measurement period or day the instrument is used.

- Surface Barrier Alpha Spectrometry Detectors. The energy calibration is often performed using an alpha source when the instrument is setup initially and when a detector has been serviced or replaced. Electronic pulsers can be used for daily checks on energy calibration. Most alpha spectra include a chemical yield tracer with a peak of known energy that can be used to verify the energy calibration during data review. Alpha spectrometers have a lower resolution than germanium detectors, and newer spectrometers are sufficiently stable to allow weekly or monthly performance checks. The frequency of performance checks should be based on the number and frequency of measurements and historical information on the stability of the instrument.

- Low-Resolution NaI(Tl) Detectors. These typically are less stable than HPGe detectors and may require more frequent quality control checks, depending on the conditions under which they are used.

For all detectors where energy calibrations are performed daily, plotting the channel numbers of peak centroids can be useful for identifying trends and determining the need for adjusting the

system. Changes in peak location may result in mis-identification of radionuclides. When this is observed, all spectra obtained since the last acceptable energy calibration check should be reviewed. If there is sufficient information within the spectrum to determine the acceptability of the energy calibration, no further action may be required for that spectrum. If the spectrum contains too few peaks of known energy, reanalysis should be initiated.

Gradual changes in peak location are not unexpected and the rate of these gradual changes can be used to establish the appropriate frequency of energy calibration checks. The acceptable limits on peak location established during the initial system setup may be used to indicate when the energy calibration needs to be readjusted.

Excursions: Changes in the energy calibration can be the result of many factors including power surges, power spikes, changes in the quality of the electrical supply, variations in ambient conditions (e.g., temperature, humidity), physical shock to the detector or associated electronics, and electronic malfunction.

Rapid changes in energy calibration are usually caused by power surges, power spikes, or physical shocks to the system. Corrective actions typically involve recalibrating the system and repeating the analysis. If changes result due to loss of cryostat vacuum, the instrument may need to be returned to the manufacturer to be refurbished or replaced.

Gradual changes in the energy calibration are usually the result of a variable or poorly conditioned power source, changes in the ambient conditions, or electronic malfunction. Corrective actions generally begin with identifying the root cause of the problem. Gradual changes that begin following relocation of the instrument are more likely to be caused by the power source or the ambient conditions. Installing a line conditioner, surge protector, and uninterrupted power supply is recommended to address problems related to the system's electrical power source. Problems with low humidity can be corrected through the use of a humidifier in dry climates or cold weather; conversely, high or variable humidity may require the use of a dehumidifier. Problems associated with fluctuations in temperature may require significant changes to the heating and cooling system for the room or building containing the instrument in order to stabilize the temperature. Gradual changes that occur following physical shocks to the system or following a rapid change in peak location with an unidentified cause are more likely to be the result of problems with the electronic equipment. In most cases the amplifier is the source of these problems, but the analog-to-digital converter, pre-amplifier, power supply voltages, and multi-channel (or single-channel) analyzer may also cause this type of problem. However, they could also be the result of crystal or detector failure. Systematic switching out of components and discussions with the instrument manufacturer will often help to identify which component may be the source of

the trouble. It may be especially difficult to identify the source of problems with new instruments in a new facility.

18.5.3.2 Peak Resolution and Tailing

Issue: The shape of the full energy peak is important for identifying radionuclides and quantifying their activity with spectrometry or spectrometry systems. Poor peak resolution and peak tailing may result in larger measurement uncertainty. If consistent problems with peak resolution are persistent , then an analytical bias most likely exists. Many factors will affect peak resolution and these are discussed below.

Discussion: Detectors with good resolution permit the identification of peaks which are close in energy. When a monoenergetic source of radiation is measured with a semiconductor, scintillation, or proportional spectrometer, the observed pulse heights have a Gaussian distribution around the most probable value (Friedlander et al., 1981). The energy resolution is usually expressed in terms of the full width at half maximum (FWHM) or the full width at tenth maximum (FWTM).

In a semiconductor detector, fluctuations in output pulse height result from the sharing of energy between ionization processes and lattice excitation (Friedlander, et al., 1981). The number of charge pairs created by radiation of a given energy will fluctuate statistically. This fluctuation occurs because the energy causes lattice vibrations in the semiconductor as well as the formation of charge pairs. This sharing of energy causes a variation in the number of charge pairs created and gives rise to the width of a measured peak. The magnitude of the statistical fluctuation is proportional to the energy of the radiation. There is also a variation in the number of charge pairs collected by a detector. This variation is accounted for by the Fano factor. Because several poorly understood factors degrade resolution in a semiconductor detector, an empirical value of the Fano factor should be used.

In a scintillation detector, the statistical fluctuations in output pulse heights arise from several sources. The conversion of energy of ionizing radiation into photons in the scintillator, the electronic emission at the photocathode, and the electron multiplication at each dynode are all subject to statistical variations. Note that the distance of the sample to the detector also impacts the resolution.

In a proportional counter, the spread in pulse heights for monoenergetic rays absorbed in the counter volume arises from statistical fluctuations in the number of ion pairs formed and the gas amplification factor (Friedlander, et al., 1981). If the gas gain is made sufficiently large, the

fluctuations in the number of ion pairs determine the resolution.

The FWHM is typically used as a measure of resolution, while the FWTM is used as a measure of tailing for the full energy peak. For Gaussian peaks with standard deviation σ, the FWHM is equal to 2.35σ. The resolution of a detector is the ratio of the FWHM to the most probable peak height. The sources of fluctuations that contribute to the standard deviation are dependent on the type of detector.

Resolution affects the ability to identify individual peaks in two ways (Gilmore and Hemingway,1995). First, it determines how close together two peaks may occur in energy and still be resolved into the two components. Second, for gamma spectrometry, when a peak of small magnitude sits on the Compton continuum of other peaks, its ability to be detected can depend on its signal-to-noise ratio. With good resolution, the available counts are distributed in fewer channels, thus those counts will be more easily identified as a peak by the spectrometry analysis software. If resolution degrades significantly the efficiency may be in error. This is especially true when the spectrum analysis involves the region of interest (ROI) concept. When the calibration is performed, the full energy peak may fit within the defined ROI limits, whereas the resolution degraded peak may have counts which fall outside them. Thus, the detector efficiency will be effectively decreased and inconsistent with the previously determined efficiency.

Tailing is another observable feature of the peak shape. Tailing is an increased number of counts in the channels on either side of the full energy peak. Tailing affects the FWTM more than the FWHM, so the ratio of FWTM to FWHM can be used as a measure of tailing. For a Gaussian distribution the ratio of FWTM to FWHM is 1.823. For most germanium detectors this ratio should not exceed 2.0. Tailing may be caused by imperfect or incomplete charge collection in some regions of the detector, escape of secondary electrons from the active region of the detector, electronic noise in the amplification and processing circuitry, loss of vacuum and escape of bremsstrahlung from the active region of the detector. Tailing may also result from the source's self-absorption for alpha emitting radionuclides.

The resolution (FWHM) is routinely calculated for gamma and alpha spectrometry peaks by the spectrum analysis software and can be monitored by observing the FWHM calculated for the check sources routinely counted. Resolution monitoring and charting is normally an integral part of a measurement quality system. Acceptance parameters may be established for resolution and incorporated in the analysis software. For alpha spectrometry, where radionuclide tracers are used for chemical yield determination, the FWHM can be monitored for each analysis, if desired. Some projects may specify FWHM limits for internal tracer peaks on each sample run.

The shape of the peak is important for quantifying the activity, and resolution is important for identifying peaks in a spectrum. The shape of the peak is also important for monitoring the performance of a detector. Germanium detectors have very good resolution on the order of 1 percent. The FWHM at specific energies is provided by the manufacturer. The FWHM should be established at several energies throughout the range being measured because the FWHM is directly proportional to the energy. These energies are usually the same as those used for checking the energy calibration and the efficiency calibration. Control limits for FWHM and the ratio of FWTM to FWHM may be developed based on statistics using multiple measurements collected over time.

The resolution of an alpha spectrum is dominated typically by self-absorption in the source. This is indicated by low energy tailing and elevated FWTM and FWHM. Most surface barrier detectors are capable of resolutions on the order of 30-40 keV for monoenergetic nuclides and 80-100 keV for unresolved multiplets. Acceptance of sample resolution is usually monitored by visual inspection of individual spectra. For well-prepared samples, the FWHM of the alpha peaks may be expected to be from 30 to 80 keV.

The resolution of scintillation detectors is not as good as the resolution of semiconductor detectors, but peak shape and tailing are just as important for analyzing samples. The FWHM should be established at several energies throughout the range being measured because the FWHM is inversely proportional to the energy. These energies are usually the same as those used for checking the energy calibration and the efficiency calibration. Control limits for FWHM and the ratio of FWTM to FWHM may be developed based on statistics using multiple measurements collected over time.

Proportional counters are not used as spectrometers in many laboratories, so it is not necessary to perform checks for resolution and peak shape.

Performance checks for resolution and tailing should be performed for all instruments used as spectrometers. These measurements are usually combined with the performance checks for energy calibration and efficiency calibration. Quality control activities should include visual inspection of all spectra to evaluate peak shape and tailing.

Control charts for FWHM and the ratio of FWTM to FWHM can be developed and used to monitor the performance of any detector used as a spectrometer. Because the concern is when the resolution degrades (i.e., the FWHM increases) or tailing becomes a problem (i.e., the ratio of FWTM to FWHM increases), control limits are necessary. Limits can be developed based on historical performance for a specific type of detector. Control charts offer a convenient method

for monitoring the results of the performance checks. As mentioned previously, the concern is associated with an increase in the FWHM or the ratio of FWTM to FWHM. This means that only an upper control limit or tolerance limit is required for the chart.

Excursions: Changes to the FWHM are associated with malfunctioning or misadjusted electronics, excessive noise or interference, or detector or source problems. Electronics problems include changes in the high voltage applied to the detector, noise (including cable noise and high voltage breakdown), and electronic drift. Electronics problems may be caused by changes in the high voltage, improper adjustment of the pole zero or baseline restorer, or drift of the amplifier gain or zero during acquisition. Source problems are usually only associated with alpha spectra and result in excessive self-absorption resulting in low-energy tailing. This can result in counts being identified with an incorrect peak. Problems that are not electronic or source related imply that the detector is malfunctioning.

Changes to the ratio of FWTM to FWHM indicate problems associated with tailing. Tailing can occur on the high- or low-energy side of the peak. High-energy tailing indicates electronics problems that may be caused by excessive activity in the sample, incorrect adjustment of the pole zero or pile-up rejector, or drift of the amplifier gain or zero while acquiring the spectrum. Low-energy tailing indicates an electronic or a source problem—a possible corrective action is to check to see if the vacuum is set properly. Table 18.4 lists common problems, the implied root cause of the problem, and possible corrective actions.

TABLE 18.4 — Root cause analysis of performance check results

Observed Problem	Implied Root Cause	Possible Corrective Actions
Efficiency changed	Unknown Electronics degradation Geometry changed Poor source Software application	Ensure the correct check source was used Check to ensure the efficiency was evaluated using the correct geometry Ensure high voltage is set properly Pulser check of electronics
Peak centroid moved	Gain changed	Check amplifier gain Check conversion gain Check stability of amplifier for gain shifts or drifting
	Offset shifted	Check zero offset Check digital offset Check stability of amplifier for gain shifts or drifting
FWHM changed	Electronics problem	Ensure high voltage is set properly Detector problem
FWTM: FWHM changed	Electronics problem	Ensure high voltage is set properly Detector problem

Observed Problem	Implied Root Cause	Possible Corrective Actions
	Source problem	Repeat sample preparation and recount Reanalyze sample Check with weightless (plated) source
No peak or broad peaks	Electronics problem	Ensure that high voltage is correct Detector problem
Low-energy tailing	Electronics problem	Ensure that high voltage is correct Check pole zero adjustment Check baseline restorer Check stability of amplifier for gain shifts or drifting Check for loss of vacuum
	Source problem	Repeat sample preparation and recount Reanalyze the sample
High-energy tailing	Electronics problem	Check pole zero adjustment Check pile-up rejector Check stability of amplifier for gain shifts or drifting
	Source problem (too much activity)	Reduce volume of sample analyzed Increase distance between the source and detector
Spectra shifted uniformly	Offset shifted	Check zero offset Check digital offset Check amplifier for zero drift
Spectra stretched or compressed	Gain changed	Check amplifier gain Check conversion gain Check amplifier for gain shifts

18.5.4 Gas Proportional Systems

18.5.4.1 Voltage Plateaus

Issue: The accuracy of the results produced by a gas proportional system can be affected if the system is not operated with its detector high voltage adjusted, such that it is on a stable portion of the operating plateau.

Discussion: The operating portion of a detector plateau is determined by counting an appropriate source at increasing increments (e.g., 50 volts) of detector high voltage. For detectors which will be used to conduct analyses for both alpha- and beta-emitting radionuclides, this should be done with both an alpha and beta source. The sources used should be similar in both geometry and energy to that of the samples to be counted in the detector.

A plot of the source count rate (ordinate) versus high voltage (abscissa) rises from the baseline to

a relatively flat plateau region, and then rises rapidly into the discharge region for both the alpha and beta determinations. From the plateau, the operating voltage is selected or verified. The operating potential is usually selected in the middle of the plateau. It remains advisable to assure that the operating point is as far as practical above the plateau knees, and in any case not less than 50 to 100 volts. Operation of the counter at the upper end of the plateau is not recommended and can result in the generation of spurious discharge counts. Modern high-voltage supplies, operating properly, experience little actual potential variance. The detector response should be checked after repairs and after a change of gas. The detector plateau should again be determined and plotted (voltage vs. count rate) after repairs, particularly to the detector unit.

The historical tracking of the establishment and maintenance of this operating parameter is recommended; it aids in determining the probable cause of quality control failure and the identification of long-term instrument deterioration. Items to be recorded include date/time, instrument detector designation, source number, check source response at the operating point, and pertinent instrument parameters, such as lower level discriminator setting, alpha discriminator setting, length of the plateau, operating high voltage setting, etc.

Excursions: Voltage changes of short- or long-term duration will affect reliability of a proportional counter. If the potential is lowered sufficiently, there is a danger of operating below the plateau knee which, in effect, reduces the efficiency and would bias the results of any sample count low. Should the voltage applied to the proportional detector be driven up to a point where the slope of the plateau is sufficiently great enough to increase the efficiency of the detector, sample counts may be biased high. A transient voltage increase of great enough magnitude could introduce spurious counts.

Shifts in the operating voltage along the plateau or length of the plateau could also result from long-term detector deterioration or electronic drift or failure.

18.5.4.2 Self-Absorption, Backscatter, and Crosstalk

Issue: The accuracy of alpha and beta activity determinations in samples with discernable solids in a gas proportional system depends in large part on the determination and maintenance of self-absorption and crosstalk curves.

Discussion: Samples counted for alpha and beta activity in a gas proportional system are typically prepared as inorganic salts, e.g., nitrates, carbonates, oxides, sulfates, or oxalates, and contain on the order of tens to hundreds of milligrams of solids when counted, which result in absorption and scattering of the particles in the sample material and mounting planchet (Chapter

16). Thus, for gas proportional systems, the detection efficiency for a given sample depends on the self-absorption occurring within each sample volume/mass. To establish the correction factor, a calibration curve is generated using a series of standards consisting of an increasing amount of solids and known amounts of radionuclide. The relative efficiency for each calibration source is plotted against the amount of solids, and these data are used to determine a sample's efficiency as a function of sample weight. The diameter and the composition of the sample planchette, not just the weight, should be identical with what was used for routine samples. This allows calculation of the corrected amount of activity regardless of the sample mass (mass/efficiency curves).

The counting of alpha and beta particles simultaneously in a proportional counter requires that an electronic discriminator be adjusted, such that pulses of heights below that represented by the discriminator are registered as betas, and those of greater heights are counted as alphas. Crosstalk occurs when alpha particles are counted in the beta channel or betas are registered as alphas. For electroplated sources, crosstalk may be as low 1 percent for betas in the alpha channel and 3 percent for alphas in the beta channel. However, this relationship is energy dependent, and care should be taken to identify samples that differ significantly from the sources used to establish the crosstalk ratio. For example, $^{90}Sr/^{90}Y$ (E_{max} 2.28 meV) is typically used as a beta source for instrument calibration. However, samples containing natural uranium in equilibrium with its progeny produce beta emissions that are considerably more energetic from the 3.28 MeV E_{max} betas of ^{214}Bi. The crosstalk ratio established with ^{90}Sr will be inadequate for such samples.

As the amount of solids in the sample increases, the alpha into beta crosstalk increases, due to the degradation of the alpha particle energy by interaction with sample material. Similarly, the beta into alpha crosstalk decreases. Thus, crosstalk should be evaluated as a function of sample weight to correct the observed relative alpha and beta counts. This is normally determined in conjunction with the self-absorption curve. To check these parameters, test samples should be prepared at the low and high ends of the calibration curve, and the limit of their acceptability should be better than 1 percent (one sigma). These checks should be performed annually at a minimum, following detector replacement or significant repair. The historical tracking of the establishment and maintenance of these operating parameters is recommended. This aids in determining the probable cause of quality control failure and the identification of long-term instrument deterioration. In addition, items to be recorded include date/time, instrument detector designation, source number, operating point, and pertinent instrument parameters, such as lower level discriminator setting, alpha discriminator setting, etc.

Excursions: Any change in the detector-source geometry or adsorption characteristics between the source and detector, can affect the self-absorption and crosstalk correction factors. For example, the replacement of a detector window with one whose density thickness is different

from the original window can necessitate the reestablishment of these parameters. Electronic drift of the alpha discriminator can also affect the crosstalk ratios.

18.5.5 Liquid Scintillation

Issue: A liquid scintillation counter is essentially a spectrometer that utilizes a multi channel analyzer to differentiate alpha or beta emission energies. These samples are subject to interferences from a variety of sources for which corrections should be made to produce useful data. A detailed discussion of liquid scintillation counting is provided in Chapter 15.

18.5.6 Summary

Table 18.5 provides some example calibration needs, performance frequency, and performance criteria, listed by detector type. Individual laboratories may be more or less stringent. These items are just presented as examples for consideration in this section. The table is presented mainly for the reader to establish their own criteria and is not intended to be a set of minimum requirements. For additional sources of information, see the calibration frequencies for several detector systems given in ASTM E181 and ANSI N42.12.

TABLE 18.5 — Instrument calibration: example frequency and performance criteria

Example Calibration Needs	Measurement Parameters	Performance Frequency	Performance Criteria
Gas Proportional System			
Initial calibration	Plateau checks as applicable	After repairs or major maintenance on control of system is re-established	Plot voltage versus counting activity to estimate proper operating voltages for both alpha and beta
	Crosstalk or sensitivity as applicable	After repairs or major maintenance on control of system is re-established	Crosstalk of alpha in beta: less than 10%; Crosstalk or sensitivity of beta in alphas: less than 1%
	Counting efficiency to calculate activity in sample	Upon incorporation of new or changes protocols	Counting uncertainty <1%; <3% uncertainty (2s) over calibration range
	Weight of solids, when mass loading is applicable, to calculate sample activity		Establish a curve for efficiency versus mass loading; <3% uncertainty (2s) over calibration range
Background counting	Count detector background using contamination-free clean planchet	One per week or batch when the system is in use	Establish a background count rate value for total alpha and beta, with N>1000

Example Calibration Needs	Measurement Parameters	Performance Frequency	Performance Criteria
Counter control or control standard	Use a source of appropriate energies	One per day when the system is in use	Control limits: three sigma or ± 3%, whichever is greater
Gamma Spectrometry			
Initial calibration	Detector energy calibration	After repairs or major maintenance if control of system cannot be re-established	Covers energy range of desired nuclides; resolution should be sufficient to separate gamma-ray lines of interest from background peaks and other interfering lines
	Counting efficiency matrix- and geometry-specific		Span energy range of nuclide of interest
Background	Counter detector background to establish background level	Minimum of every week or after analytical run, whichever is longer	
Counter control or control standard	Multi energy source covering the general energy calibration range	One per week or after analytical run, whichever is longer	Control limits: three sigma or ± 3%, whichever is greater
Alpha Spectrometry			
Initial calibration	Energy calibration	After repairs or major maintenance if control of system cannot be re-established	No specific criteria, pending on total channel and range of energy spectrum of desired nuclides
	Counting efficiency matrix- and geometry-specific		Span energy range of nuclide of interest
Background	Counter detector background to establish background level	Minimum of every other week or after analytical run, whichever is longer	
Counter control or control standard	At least two isotopes Monitor peak location, resolution and efficiency (where counting efficiency is an analytical requirement).	One per week or after analytical run, whichever is longer	Control limits: three sigma or ± 3%, whichever is greater
Liquid Scintillation			
Initial Calibration	Dark blank to check photomultiplier tube	After mechanical or electronic repairs	Check against manufacturer's specifications
Calibration	External (instrumental) calibration	After repairs or major maintenance if control of system cannot be re-established	Check against manufacturer's specifications

Example Calibration Needs	Measurement Parameters	Performance Frequency	Performance Criteria
Method Calibration (Determining quenching)	Quench curve (at least five points)	If matrix or cocktail changes	
	Internal standard	Add to each sample type	
Background	Counter detector background	One per day or analytical batch when the system is in use	
Counter control or control standard		One per day or batch when system is in use	Control limits: three sigma or ± 3%, whichever is greater
Batch-approach calibration (Alternative approach)	Minimum two matrix-matched standards and blanks	One per batch	Counting efficiency control limits: three sigma or ± 5%, whichever is greater

Sources: ASTM E181; ANSI N42.12.

18.5.7 Non-Nuclear Instrumentation

Radioactivity and radionuclide measurement techniques also employ the use of non-nuclear instrumentation such as mass spectrometry, fluorimetry, phosphorimetry, and fission tract. Although these instruments are not covered in MARLAP, analysts can apply many of the laboratory QC techniques discussed in Sections 18.3, 18.4, and 18.6 because they are basic to any laboratory method. A quality program using statistically based control charts of the performance indicators will identify out of control situations, assist in improving laboratory performance and aid in identifying the causes of trends and biases for any laboratory method. Analysts also need to consider detection capabilities, radionuclide secular equilibrium, half-life, interferences, and blind samples when using non-nuclear instrumentation.

18.6 Related Concerns

18.6.1 Detection Capability

Issue: The *detection capability* of an analytical procedure is its ability to distinguish small amounts of analyte from zero (Chapter 19). The detection capability of a procedure can be estimated nominally and will depend on many factors.

Discussion: In radioanalysis, the most commonly used measure of detection capability is the minimum detectable concentration (Chapter 19). The MDC is defined as the smallest concentration of an analyte that has a specified probability of detection, typically 95 percent. The MDC is usually estimated as a nominal scoping performance measure of an analytical procedure, but a

sample-specific version is reported routinely by many laboratories.

Detection capability is affected by many factors, including counting times, instrument background levels, aliquant volume, yield, decay times, and interferences. The nominal MDC is presumably based on conservative assumptions about these factors, but measurement conditions vary. The sample-specific MDC is calculated using the actual measured values of all these factors. A high MDC by itself does not indicate that a sample result is invalid or that it cannot be used for its intended purpose. However, if an analysis fails to detect the analyte of interest and the sample-specific MDC is greater than a detection limit required by contract or other agreement, it may be necessary to reanalyze the sample in a way that reduces the MDC. Such decisions should be made case-by-case, since it is not always cost-effective or even possible to reanalyze a sample, or it may not be feasible to achieve the desired MDC.

Excursions: A high sample-specific MDC can be caused by many factors, including:

- Small sample aliquant;
- Low chemical/tracer yield;
- Short counting times;
- Long decay/short ingrowth time;
- High background or blank value; and
- Low counting efficiency or sample self-attenuation.

18.6.2 Secular Equilibrium

Issue: It is sometimes necessary to ensure that target radionuclides are in secular equilibrium with their progeny, or to establish and correct for disequilibrium conditions. This is particularly applicable for protocols that involve the chemical separation of long-lived radionuclides from their progeny. This is also applicable for nondestructive assays like gamma spectrometry where photon emission from progeny is used to determine the concentration of the non-gamma ray emitting parent.

Discussion: Some radionuclides that have long physical half-lives decay to species whose half-lives are shorter by several orders of magnitude. Following chemical separation of the parent, the progeny can "grow in" within a time frame relevant to analysis and provide measurable radioactive disintegration which should be considered in the analytical method. The condition where the parent and progeny radionuclide are equal in activity is called "secular equilibrium." An example is ^{226}Ra, a common, naturally occurring radionuclide in the uranium series with a half-life of about 1,600 years. ^{226}Ra is found in water and soil, typically in secular equilibrium with a

series of shorter-lived radionuclides that begins with the 3.8-day-half-life ^{222}Ra and ends with stable lead. As soon as ^{226}Ra is chemically separated from its progeny in an analytical procedure via coprecipitation with barium sulfate, its progeny begin to reaccumulate. The progeny exhibit a variety of alpha, beta and gamma emissions, some of which will be detected when the precipitate is counted. The activity due to the ingrowth of radon progeny should be considered when evaluating the counting data (Kirby, 1954). If counting is performed soon after chemical separation, secular equilibrium will be substantially incomplete and a sample-specific correction factor should be calculated and applied. In some cases, it may be necessary to derive correction factors for radioactive ingrowth and decay during the time the sample is counting. These factors are radionuclide specific, and should be evaluated for each analytical method.

Secular equilibrium concerns also apply to non destructive assays, particularly for uranium and thorium series radionuclides. Important radionuclides in these series (e.g., ^{238}U and ^{232}Th) have photon emissions that are weak or otherwise difficult to measure, while their shorter-lived primary, secondary or tertiary progeny are easily measured. This allows for the parents to be quantified indirectly, i.e., their concentration is determined by measuring their progeny and accounting for the amount of parent-progeny equilibrium. The amount of parent-progeny secular equilibrium is fundamental to these analyses, and data should be scrutinized to insure that the amount is valid.

When several radionuclides from one decay chain are measured in a sample, observed activity ratios can be compared to those predicted by decay and ingrowth calculations, the history of the sample and other information. For example, undisturbed soil typically contains natural uranium with approximately equal activities of ^{238}U and ^{234}U, while water samples often have very different ^{238}U/^{234}U ratio. Data from ores or materials involved in processing that could disrupt naturally occurring relationships require close attention in this regard.

All calculational protocols (electronic and manual) should be evaluated to determine if there is bias with respect to correction factors related to equilibrium concerns. This includes a check of all constants used to derive such correction factors, as well as the use of input data that unambiguously state the time of all pertinent events (chemical separation and sample counting). The analyst should ensure that samples requiring progeny ingrowth are held for sufficient time before counting to establish secular equilibrium. Limits for minimum ingrowth and maximum decay times should be established for all analytical methods where they are pertinent. For ingrowth, the limits should reflect the minimum time required to ensure that the radionuclide(s) of interest has accumulated sufficiently to not adversely affect the detection limit or uncertainty. Conversely, the time for radioactive decay of the radionuclides of interest should be limited such that the decay factor does not elevate the MDC or adversely affect the measurement uncertainty. These will

vary depending on the radionuclide(s) and analytical method.

Excursions: Samples where equilibrium is incorrectly assumed or calculated will produce data that do not represent the true sample concentrations. It is difficult to detect errors in equilibrium assumptions or calculations. Frequently, it takes anomalous or unanticipated results to identify these errors. In these cases, analysts need to know the sample history or characteristics before equilibrium errors can be identified and corrected. Some samples may not be amenable to nondestructive assays because their equilibrium status cannot be determined; in such cases, other analytical methods are indicated.

Examples:

Isotopic Distribution – Natural, Enriched and Depleted Uranium: Isotopic distribution is particularly important with respect to uranium, an element that is ubiquitous in nature in soils and also a contaminant in many site cleanups. The three predominant uranium isotopes of interest are ^{238}U, ^{234}U, and ^{235}U, which constitute 99.2745, 0.0055, and 0.72 atom percent, respectively, of "natural" uranium[3], i.e., uranium as found in nature (General Electric, 1984). However, human activities related to uranium typically involve changing the ratio of natural uranium by separating the more readily fissionable ^{235}U from natural uranium to produce material "enriched" in ^{235}U, for use in fuel cycle and nuclear weapons related activities. Typical ^{235}U enrichments range from 2 percent for reactor fuels to greater than 90 percent ^{235}U for weapons. The enrichment process also produces material that is "depleted" in ^{235}U, i.e., the uranium from which the ^{235}U was taken.[4] While the ^{235}U concentrations of depleted uranium are reduced relative to natural ores, they still can be measured by several assay techniques. This gives rise to uranium with three distinct distributions of ^{238}U, ^{235}U, and ^{234}U, referred to as "natural," "enriched," and "depleted" uranium. Because ^{238}U, ^{235}U, and ^{234}U are alpha emitters with considerably different physical half-lives and specific activity, a measurement of a sample's total uranium alpha activity cannot be used to quantify the sample's isotopic composition or uranium mass without knowing if the uranium is natural or has been enriched or depleted in ^{235}U. However, if this information is known, measurement and distribution of the sample's uranium alpha activity can be used to infer values for a sample's uranium mass and for the activities of the isotopes ^{238}U, ^{235}U, and ^{234}U. This ratio can be determined directly or empirically using mass or alpha spectrometry, techniques which are

[3] The "natural abundance" of ^{235}U of 0.72 atom percent is a commonly accepted average. Actual values from specific ore samples vary.

[4] Enriched and depleted refer primarily to ^{235}U.

time and cost intensive, but which provide the material's definitive isotopic distribution. It is often practical to perform mass or alpha spectrometry on representative samples from a site to establish the material's isotopic distribution, assuming all samples from a given area are comparable in this respect. Once established, this ratio can be applied to measurements of uranium alpha activity to derive activity concentrations for ^{238}U, ^{234}U, and ^{235}U data.

18.6.3 Half-Life

Issue: Radionuclides with short half-lives relative to the time frame of the analysis may decay significantly from the time of sample collection or chemical separation to counting. In some cases, this decay will cause the ingrowth of other short-lived radionuclides. In both instances, sample-specific factors should be applied to correct the sample's observed counting/disintegration rate. Also, determination of half-life could indicate sample purity. If radioactive impurities are not appropriately corrected, analytical errors will occur. Consecutive counting of the sample may confirm the radionuclide impurity by analyzing the decay rate between counting events.

Discussion: When assaying for short-lived radionuclides, data should be corrected for decay over the time period between sample collection and counting. For example, operating power reactors routinely assay environmental samples for ^{131}I, a fission product with about an eight-day half-life. Samples may be counted for several days up to two weeks, during which time their ^{131}I concentration is decreasing via radioactive decay. Using the eight-day half-life, the counting data should be decay-corrected to the time of collection in the field. If desired, environmental samples can be decay-corrected to a time other than sample collection.

Half-life considerations also apply to radionuclide ingrowth. Certain radionuclides are assayed by an initial chemical separation which begins a period over which their direct progeny are allowed to come to secular equilibrium; this is followed by chemical separation, purification and counting of the progeny. After counting, the degree of the progeny's ingrowth is calculated, based on the radionuclides' half-lives and the elapsed time between separation and counting. Allowance should also be made for the progeny's decay from separation to counting and for decay that occurred while counting, if applicable. Two examples are the beta emitting radionuclides ^{228}Ra and ^{90}Sr: they are quantified by measuring the direct progeny of each, ^{228}Ac and ^{90}Y, respectively. For airborne concentrations of ^{222}Rn, sample collection and analytical methods should incorporate concerns related to the short-lived progeny of other radon species, such as ^{220}Rn. Other half-life related considerations apply to alpha spectrometry when assaying samples for uranium and thorium chain radionuclides. Samples that have been allowed to sit for several weeks may accumulate short-lived radionuclides that have alpha emissions whose energies are in close proximity to target radionuclides. These can interfere with quantitative analyses of the target

radionuclides. Chemical yield tracers used in alpha spectrometry, such as ^{234}Th and ^{232}U, can cause this effect due to their short-lived progeny and all chemical yield tracers should be scrutinized for this potential prior to their use in analytical methods. Radionuclide specific limits for minimum ingrowth and maximum decay times should be established for all analytical methods where they are pertinent. These should be based on limiting the adverse effect of such calculations on the detection limit and measurement uncertainty. All analytical methods involving computational corrections for radioactive decay of the target species should be evaluated relative to half-life and secular equilibrium related concerns. This evaluation should be incorporated in the routine data review process that is performed on all analytical results.

A good source for radionuclide half-lives and other nuclear data can be found at the Brookhaven National Laboratory's National Nuclear Data Center (http://www.nndc.bnl.gov/nndc/nudat/). Using this data source will ensure consistency within and among laboratories, and will provide analysts with the current values.

Excursions: Samples that are assayed by "non destructive" techniques like gamma spectrometry may provide indications of potential complications due to half-life related considerations. Because the assay provides information on photon emitting radionuclides in the sample, the analyst can develop appropriate corrections for half-life related phenomena. However, non-spectrometric techniques like gas flow proportional counting are essentially gross counting procedures that record all events without any indication of their origin. Therefore, these data should be evaluated to ensure they are free from half-life related considerations.

Samples with short-lived radionuclide concentrations at or near environmental background will experience elevated detection limits and increased measurement uncertainty if there is excessive elapsed time between sample collection and counting. Because there is an additional correction factor in the algorithms for these samples (decay factor), they are more susceptible to measurement uncertainty than longer-lived radionuclides.

18.6.4 Interferences

Issue: Chemical or radionuclide interferences can produce erroneous results or increased measurement uncertainty.

Discussion: Analytical samples, particularly environmental samples, are often chemically complex. This complexity may include chemical constituents or other physical aspects that interfere with an analytical method to the point that they require modification of the method. Examples of modifications include limiting the size of the sample aliquant, quantifying

interfering compounds through other analyses (radiometric and non-radiometric) and changing time periods to allow adequate ingrowth of target radionuclides or decay of interferences.

A common example is groundwater or well water that contains high concentrations of salts or dissolved solids, so that screening for gross alpha activity produces erratic or anomalous results. For such samples, it may be necessary to limit the aliquant volume with the resulting increase in detection limit and measurement uncertainty. There is a concentration at which this procedure cannot overcome the interferences and should not be used.

Samples that contain natural concentrations of stable or unstable compounds that an analytical procedure adds to the sample for a specific purpose (carrier or tracer) may also be problematic because the sample's concentration interferes with the analysis. Because barium is used as a carrier, water samples that contain high concentration of barium may provide inaccurate carrier yields when screened for alpha-emitting radium isotopes. Quantifying the sample's barium content prospectively via a non-radiometric technique (e.g., atomic absorption) would be required to correct for this interference. With respect to unstable compounds, two examples are provided. The first involves the radiochemical procedure for determining ^{228}Ra in drinking water that separates radium via coprecipitation with barium sulfate. The precipitate is allowed to come to equilibrium with its direct progeny ^{228}Ac, which is separated via co-precipitation with yttrium oxalate, purified, mounted and counted. The yttrium precipitate also carries ^{90}Y, the direct progeny of ^{90}Sr, a fission product often found in environmental samples as a result of atmospheric weapons testing and nuclear fuel cycle activities. Samples assayed for ^{228}Ra may contain measurable amounts of ^{90}Sr that require corrections based on differences in half-life (^{228}Ac with a 6-hour half-life versus ^{90}Y with a half-life of about 64 hours) or other parameters. The second example involves alpha spectrometry procedures that use tracers to determine chemical yield. For example, ^{234}Th is used as a chemical yield tracer for isotopic thorium analyses. The approach assumes that the sample's inherent concentration of the tracer radionuclide is insignificant such that it will not interfere with the tracer's ability to accurately represent the sample's chemical recovery. Samples that contain measurable amounts of these radionuclides may produce excessive interference and may not be amenable to this procedure.

Alpha spectra should be checked for radionuclide interferences, e.g. look for ^{238}U peak in a Pu spectra. If the ^{238}U peak is present, ^{234}U might be an interference in the ^{239}Pu and ^{240}Pu determinations. Data can be corrected or the sample may require reanalysis.

Each analytical method should be evaluated with respect to interferences, when its use is proposed or at least prior to their implementation in the laboratory. Such evaluations can be based on available information and, if properly documented, can serve as the basis for developing

the range of applicability, which becomes an integral part of the protocol. Evaluating performance indicators aids in the identification of samples that have interferences. All performance criteria would be protocol specific, and have clearly established acceptance ranges that incorporate the potential interferences discussed above.

Excursions: Interfering elements can affect measurement results in several ways. For example, large amounts of non-analyte elements may overload ion exchange resins, affecting the resin's ability to collect all of the analyte. In addition, spiking elements, already in the sample prior to preparation, may cause matrix spike results to exceed acceptance limits.

Carrier/tracer yields exhibiting gradual changes that appear to be correlated with a batch or group of samples from the same sampling location may indicate potentially interfering conditions. A significant decrease in the carrier/tracer recovery may indicate that the analytical method is not functioning as planned. Yields that are significantly low or in excess of 100 percent may be caused by competing reactions within the sample matrix, or by the presence of inherent concentrations of carrier/tracer within the sample.

For screening analyses, e.g., gross alpha or beta, large changes in counting efficiencies or erratic counting data can reflect the presence of salts. Samples of this type are hydroscopic, and continue to gain weight following preparation in planchettes as they absorb moisture from the air. These changes could be detected by reweighing the planchettes directly prior to counting. These samples can be converted to oxides by carefully holding them over the open flame of a laboratory burner; however, this will cause losses of volatile radionuclides, predominantly ^{210}Po and ^{137}Cs, which have alpha and beta emissions, respectively. An alternative approach is to thoroughly dry each planchette, record the weight and count it immediately, followed by a post-counting weighing to ensure that the weight did not change significantly over the measurement period. This approach may not be practical for all laboratories.

18.6.5 Negative Results

Issue: When an instrument background measurement is subtracted from a measurement of a low-activity sample, it is possible to obtain a net activity value less than zero.

Discussion: Many factors influence the evaluation of negative results. The simplest case occurs when the background measurement is unbiased and both the gross counts and background counts are high enough that the distribution of the net count rate is approximately normal. In this case, normal statistics can be used to determine whether a negative result indicates a problem. For example, if a sample contains zero activity, there is a very small probability of obtaining a net

count rate more than two-and-a-half or three standard deviations below zero Since the combined standard uncertainty is an estimate of the standard deviation, a result that is less than zero by more than three times its combined standard uncertainty should be investigated. In fact, if a blank sample is analyzed using an unbiased measurement process, negative results can be expected about 50 percent of the time. As long as the magnitudes of negative values are comparable to the estimated measurement uncertainties and there is no discernible negative bias in a set of measurements, negative results should be accepted as legitimate data and their uncertainty should be assessed. On the other hand, if a sample activity value is far below zero, there may be a reason to investigate the result. A large percentage of negative results may also indicate a problem, even if all of the results are near zero. When instrument backgrounds are extremely low, statistics based on a normal distribution may not be appropriate (Chapter 19).

A preponderance of results that are negative, even if they are close to zero, indicates either a systematic error or correlations between the results. If the results are measured independently, a pattern of negative results indicates a bias, which requires investigation.

Excursions: Negative results occur routinely when samples with low levels of activity are analyzed, but a result should seldom be more than a few standard deviations below zero. Possible causes for extremely negative results or for an excessive number of negative values include:

- Instrument failure (low sample counts or high blank counts);
- Positive bias in the background or reagent blank measurement;
- Overestimation of interferences;
- Data transcription error; or
- Calculation error.

18.6.6 Blind Samples

Issue: The performance of the analytical method should be assessed independently on a regular basis. This assessment is achieved through the use of blind samples that provide an objective means of evaluating the laboratory's performance for specific analytes and matrices. Blind samples can be internal or external, and either single or double. External blind PE samples are used for QA purposes and also can provide information that is useful to laboratory QC.

Discussion: A blind sample is a sample whose concentration is not known to the analyst, and whose purpose is to assess analytical performance. Regardless of their nature, blind samples are effective only when their contents are unknown to the analysts. The preparation of all blind and other performance assessment samples is usually designated as a QA function. The QA staff

functions independently from personnel responsible for sample processing and analysis. Blind samples consist of a matrix routinely processed by the laboratory that contains a known amount of one or more analytes (radionuclides). A blind also may take the form of a replicate sample that is submitted for analysis such that its composition and origin are unknown to the analyst. These can be split samples (if run in the same batch) or spiked samples, and are prepared and submitted by an independent group either within the organization (internal), or from an independent organization (external). Performance on blind samples should be an integral part of the laboratory's quality system, which includes routine evaluation of them against specific performance criteria. For example, analysis of blind samples should be evaluated for relevant performance indicators. Data that fall outside an acceptance criterion may indicate loss of control in sample chemical processing, radiometric determination (counting) or other aspects of the analytical process. The ability to prepare blind samples depends fundamentally on the ability to obtain the appropriate combination of matrix with a radionuclide of a well-known concentration, ideally traceable to NIST or other appropriate certifying body. Also important are the expertise and experience of the preparer of the blind samples, proven and verified methodologies used for the blind samples, and detailed documentation. The use of blind samples assumes that their physical, chemical and radiological nature are compatible with the analytical methods employed at the laboratory.

When the analyst is aware that the sample is a blind sample but does not know the concentration, these samples are called single blinds. In the case of replicates , the analyst is not aware that two samples are the same; for spiked samples, the analyst may know what analytes the blind sample contains, but not the analyte's concentration. Single blinds and other internal samples of this type are generally prepared by an organization's QA personnel that are independent of the samples' analyses. External single blind samples are available and can be obtained from several sources.

A double blind sample is the same as a single blind except that it is submitted for analysis as a routine sample. The sample should be identical in appearance to a routine sample, and the analyst is not forewarned of the analytes in the sample. In general, a double blind is thought to be a more rigorous indication of the laboratory's performance, since analysts and other laboratory personnel may take special precautions when analyzing known PT samples, in anticipation of the greater scrutiny associated with such samples. This should not happen with double blind samples, since there should be no way to distinguish them from routine samples. However, true double blind samples are difficult to prepare.

INTERNAL BLIND SAMPLES. Internal blind samples are prepared by the laboratory's QA personnel. Internal blind samples assess several aspects of the analytical process. They allow the laboratory to demonstrate that it can successfully process routine samples for a specific

analysis; in other words, they get a measured result within accepted limits. They provide an auditable, empirical record against specific quality performance criteria. They also demonstrate the efficacy of analytical methods and areas in need of adjustment. Double blind samples can pose logistical problems. It may be difficult to prepare internal double blind samples and submit them to the laboratory for analysis successfully disguised as routine samples. Evaluation criteria should be established to identify when conditions are out of acceptance limits.

EXTERNAL BLIND SAMPLES. External blind samples are those prepared by an organization outside that laboratory. This may be helpful with respect to ensuring that the analyte concentrations are truly unknown to the analyst; external blinds may offer a greater variety of matrices and analytes than can easily be produced within the laboratory and augment the laboratory's internal quality control program. Alternatively, if external blinds are not appropriate to the laboratory's programs, they will be of limited utility.

If differences between observed and known values typically arise, these should be investigated thoroughly, as they indicate areas where important details of the analytical process may have been overlooked. Often a laboratory's observed values agree with the known value within acceptable tolerances, but are biased high or low. Careful documentation of the laboratory's performance in this regard can assist in characterizing the fluctuations of a measurement system or analytical method. Like other performance indicators, large or sudden changes in bias require scrutiny.

Blind samples should be an integral part of the laboratory's quality control program and they should be processed according to a predetermined schedule. Important sources of external blind samples include the NIST Radiochemistry Intercomparison Program (NRIP), National Voluntary Accreditation Program (NVLAP/EPA), Food and Drug Administration, DOE Lab Accreditation Program (DOELAP), Quality Assessment Program (DOE QAP), and Multi-Analyte Performance Evaluation Program (DOE MAPEP).

Excursions: The excursions typically encountered with analytical methods for specific parameters (carrier/tracer recovery, lack of precision, elevated backgrounds, etc.) apply to blind samples as well. Additionally, instances where the analysis of external blinds produces values that do not agree with the known values, may indicate that instrument calibrations or other correction factors require reevaluation. Problems revealed by the analysis of blind blank samples can indicate a problem (e.g., bias, blunder) within the laboratory, or conditions where the current protocol is inadequate. Excursions discovered while analyzing samples from external PE programs should be addressed.

18.6.7 Calibration of Apparatus Used for Weight and Volume Measurements

Issue: Fundamental to all quantitative analysis is the use of the proper weights and volumes. Analysts should perform careful gravimetric and volumetric measurements (especially in the preparation of calibration solutions, test sources, and reagents) in order to achieve the desired levels of precision and bias in each analytical method. Therefore, laboratory balances and volumetric glassware and equipment should be calibrated and checked periodically to maintain the desired method performance levels. This section discusses the calibrations of laboratory balances and volumetric glassware and equipment.

Discussion: Laboratory balances should be periodically calibrated and checked. Most balances are typically calibrated and certified by the manufacturer once a year. These calibrations are performed to achieve the manufacturer's specified tolerances for each balance. A calibration certificate is supplied to the laboratory. In addition to this yearly calibration, daily calibration checks should be performed by the laboratory. Some laboratories check the balances once a day or at the time of each use. Any balance failing the daily calibration check should be taken out of service. Ordinarily, ASTM E617 Class 1 or 2 weights are used to perform the daily calibration check, depending on application. Over time, daily wear and tear on the weights can affect calibration, so it is a good idea to get them periodically re-certified or to purchase new weights.

Volumetric glassware and equipment, especially those used in the preparation of instrument calibration solutions and laboratory control samples, should be calibrated to the desired level of accuracy. Calibration can either be performed by the manufacturer of the equipment or by laboratory personnel. Calibration certificates for volumetric pipets and flasks are provided by the manufacturer at the time of purchase. Borosilicate and pyrex volumetric glassware will hold its calibration indefinitely provided that it is not exposed to hydrofluoric acid, hot phosphoric acid or strong alkalis, and that it is not heated above 150 °C when drying. Any glass volumetric pipet with a damaged tip should be discarded or re-calibrated. The manufacturer of volumetric automatic pipetting equipment calibrates the equipment and provides a certificate at the time of purchase. The re-calibration of automatic equipment should be performed annually and can be performed by the manufacturer, calibration specialty companies, or in-house laboratory personnel. Outside calibration services should provide a calibration certificate.

Laboratory personnel can calibrate and check volumetric apparatus using procedures like those specified in ASTM E542. Typically calibrations use volumes of water and are gravimetrically based. Volumes are corrected for temperature and atmospheric pressure and require thoroughly cleaned glassware, standard procedures for setting and reading the water meniscus, and accurate balances and thermometers.

Volumetric glassware is calibrated either "to contain" (TC) or "to deliver" (TD). Glassware designated as "to contain" requires the complete emptying of the vessel to yield the specified volume. "To deliver" glassware does not require complete emptying. Specified volumes for this type of apparatus do not include the residual left from surface adhesion and capillary action. TD glassware will perform with accuracy only when the inner surface is so scrupulously clean that the water wets it immediately and forms a uniform film when emptying.

18.7 References

18.7.1 Cited Sources

American National Standards Institute/International Standards Organization/American Society for Quality Control (ANSI/ISO/ASQC) A3534-2. *Statistics–Vocabulary and Symbols–Statistical Quality Control*.

American National Standards Institute/American Society for Quality Control (ANSI/ASQC) E4. 1994. *Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs.*

American National Standards Institute (ANSI) N1.1. *American Nuclear Standard Glossary of Terms in Nuclear Science and Technology*, 1976.

American National Standards Institute (ANSI) N15.37. *Guide to the Automation of Nondestructive Assay Systems for Nuclear Material Control*. 1981.

American National Standards Institute (ANSI) N42.12. American National Standard. *Calibration and Usage of Thallium-Activated Sodium Iodide Detector Systems for Assay of Radionuclides*.

American National Standard Institute (ANSI) N42.23. *Measurement and Associated Instrumentation Quality Assurance for Radioassay Laboratories.* 1996.

American Society for Testing and Materials (ASTM) D3648, *Standard Practices for the Measurement of Radioactivity*, 1995.

American Society for Testing and Materials (ASTM) D6299, *Standard Practice for Applying Statistical Quality Assurance Techniques to Evaluate Analytical Measurement System Performance,* 2000

American Society for Testing and Materials (ASTM) E542, *Standard Practice for Calibration of Laboratory Volumetric Apparatus*, 2000.

American Society for Testing and Materials (ASTM) E617, *Standard Specification for Laboratory Weights And Precision Mass Standards*, 1997.

American Society for Testing and Materials (ASTM) E181, Standard Test Methods for *Detector Calibration and Analysis of Radionuclides*.

American Society for Testing and Materials (ASTM) E882, Standard Guide for *Accountability and Quality Control in the Chemical Analysis Laboratory.*

American Society for Testing and Materials (ASTM) MNL 7, *Manual on Presentation of Data and Control Chart Analysis* ASTM Manual Series, 6th Edition, 1990.

Friedlander, G., Kennedy, J.W., Macias, E.S., and Miller, J.N. 1981. *Nuclear and Radiochemistry*. 3rd Edition, John Wiley and Sons, New York.

General Electric Company. 1984. Chart of the Nuclides, Thirteenth Edition.

Gilmore, G. and Hemingway, J.D. 1995. *Practical Gamma-Ray Spectrometry*. Wiley, Chichester, England.

International Standards Organization (ISO) 5725-1. *Accuracy (Trueness and Precision) of Measurement Methods and Results—Part 1: General Principles and Definitions.*

International Standards Organization (ISO) 7870. *Control Charts – General Guide and Introduction.*

International Standards Organization (ISO) 7873. *Control Charts for Arithmetic Average With Warning Limits*.

International Standards Organization (ISO) 7966. *Acceptance Control Charts.*

International Standards Organization (ISO) 8258. *Shewhart Control Charts*.

International Standards Organization/International Electrotechnical Commission (ISO/IEC) 17025. *General Requirements for the Competence of Testing and Calibration Laboratories*.

December 1999, 26 pp.

Kirby, H.W. 1954. Decay and Growth Tables for the Naturally Occurring Radioactive Series. *Anal. Chem*. 26:6, p. 1063-1071.

Lin, Z., K. G. W. Inn, and J. J. Fiilben. 2001. An alternative statistical approach for interlaboratory comparison data evaluation. *Journal of Radioanalytical and Nuclear Chemistry*, 248:1, 163-173.

National Council on Radiation Protection and Measurements (NCRP) 58: A Handbook of Radioactivity Measurement Procedures, Second Edition. Bethesda, MD. February 1985. (Supersedes First Edition, November 1978.)

National Environmental Laboratory Accreditation Conference (NELAC). 2000. *Quality Systems* Appendix D, *Essential Quality Control Requirements*. Revision 14. June 29. Available at http://www.epa.gov/ttn/nelac/2000standards.html.

National Bureau of Standards (NBS). 1964. Handbook of Mathematical Functions. M. Abramowitz and Stegun, I., Editors.

U.S. Environmental Protection Agency (EPA). 1977. *Handbook for Analytical Quality Control in Radioanalytical Laboratories.* EPA-600-7-77-088.

U.S. Environmental Protection Agency (EPA). 1980. *Prescribed Procedures for Measurement of Radioactivity in Drinking Water—Procedure 904.0, Determination of Radium-228 in Drinking Water*. EPA 600-4-80-032.

U.S. Environmental Protection Agency (EPA). 1980. *Prescribed Procedures for Measurement of Radioactivity in Drinking Water—Procedure 908.1 for Total Uranium in Drinking Water*. EPA 600-4-80-032.

18.7.2 Other Sources

American National Standards Institute (ANSI) N42.22. American National Standard. *Traceability of Radioactive Sources to the National Institute of Standards and Technology (NIST) and Associated Instrument Quality Control*.

Chase, G.D. and Rabinowitz, J.L. 1969. *Principles of Radioisotope Methodology*. 3rd Edition,

1716 Burgess Publishing Co., Minneapolis, MN.

1717 Kanipe, L.G. 1977. *Handbook for Analytical Quality Control in Radioanalytical Laboratories*.
1718 EPA-600/7-77-088.

1719 U.S. Environmental Protection Agency (EPA). 1995. *Guidance for the Preparation of Standard*
1720 *Operating Procedures (SOPs) for Quality-related Documents*. QA/G-6. EPA 600-R-96-027.
1721 Available at http://www.epa.gov/oerrpage/superfund/programs/clp/download/epaqag6.pdf.

1722 Zeigler, L.H. and Hunt, H.M. 1977. *Quality Control for Environmental Measurements Using*
1723 *Gamma-Ray Spectrometry*.EPA 600-7-77-144.

Attachment 18A: Control Charts

18A.1 Introduction

This attachment provides statistical details to augment Section 18.3.2. The term "statistical quality control" refers to QC based on statistical principles. Generally, statistical QC in the laboratory applies the principles of hypothesis testing, with varying degrees of rigor, to make inferences about a measurement system or process. The primary tool for statistical QC is the control chart.

The most important purpose for statistical QC in the laboratory is to ensure that measurement uncertainties are properly estimated. The uncertainty estimate that accompanies a measured value may be misleading unless the measurement process is in a state of *statistical control*. Statistical control implies that the distribution of measured results is stable and predictable. It exists when all the observed variability in the process is the result of random causes that are inherent in the process. The existence of variability due to "assignable" causes, including instrumental and procedural failures and human blunders, which are not inherent in the process, implies that the process is unpredictable and hence "out of control."

Statistical QC procedures are designed to detect variability due to assignable causes. When such variability is detected, specific corrective action is required to determine the cause and bring the measurement process back into a state of statistical control. Laboratory QC procedures should be strict enough to detect variations in the measurement system that could have a significant impact on measurement uncertainties.

Statistical QC also may be used in the laboratory to monitor method performance parameters, such as chemical yield, to ensure that the measurement system is performing as expected. However, the need for corrective action in the case of a low yield may not be as urgent as in the case of a malfunctioning radiation counter, since the latter is much more likely to cause underestimation of measurement uncertainties.

The following sections describe the various types of control charts introduced in Section 18.3.2, including the X chart, $\overline{X}$ chart, R chart, and variants of the c chart and u chart for Poisson data.

18A.2 X Charts

Procedure 18.1, shown below, may be used to determine the central line, control limits, and warning limits for an X chart. Ideally, the data distribution should be approximately normal,

1754 although the *X* chart is often used with other types of distributions. (The data may be tested for
1755 normality using the procedure described in Attachment 19F.)

1756 In order to use Procedure 18.1, an unbiased estimate of the standard deviation of the measured
1757 values $X_1, X_2, \ldots, X_n$ is required. Although the experimental variance s^2 of the data is an unbiased
1758 estimate of the true variance σ^2, taking the square root of s^2 generates a bias . The experimental
1759 standard deviation s is given by the equation

1760
$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2} \qquad (1)$$

1761 If the data are (approximately) normally distributed, s should then be divided by the value of c_4
1762 shown in Table 18A-1 below for the number of degrees of freedom $\nu = n - 1$. Thus, σ is esti-
1763 mated by s / c_4. The factor c_4 is equal to

1764
$$c_4 = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\sqrt{\frac{2}{n-1}} \qquad (2)$$

1765 where Γ denotes the *gamma function* (NBS 1964), but it is well approximated by $c_4 \approx \frac{4n-4}{4n-3}$. For
1766 large n the value of c_4 is approximately 1.

TABLE 18A-1 — Bias-correction factor for the experimental standard deviation

$\nu = n - 1$	c_4	ν	c_4	ν	c_4	ν	c_4
1	0.79788	11	0.97756	21	0.98817	31	0.99197
2	0.88623	12	0.97941	22	0.98870	32	0.99222
3	0.92132	13	0.98097	23	0.98919	33	0.99245
4	0.93999	14	0.98232	24	0.98964	34	0.99268
5	0.95153	15	0.98348	25	0.99005	35	0.99288
6	0.95937	16	0.98451	26	0.99043	36	0.99308
7	0.96503	17	0.98541	27	0.99079	37	0.99327
8	0.96931	18	0.98621	28	0.99111	38	0.99344
9	0.97266	19	0.98693	29	0.99142	39	0.99361
10	0.97535	20	0.98758	30	0.99170	40	0.99377

An alternative method of estimating the standard deviation is based on the average value of the *moving range* (ASTM D6299, ASTM E882). The moving range (MR) is the absolute value of the difference between consecutive measured values X_i and X_{i+1}. If the data are normally distributed, the expected value of the moving range is

$$\frac{2\sigma}{\sqrt{\pi}} \approx 1.128\,\sigma \tag{3}$$

which may be estimated by

$$\overline{\mathrm{MR}} = \frac{1}{n-1}\sum_{i=1}^{n-1}\left|X_{i+1} - X_i\right| \tag{4}$$

So, σ is estimated by $\overline{\mathrm{MR}} / 1.128$. The moving-range estimate of σ may be preferred because it is less sensitive to outliers in the data. Furthermore, when consecutive values of X_i are correlated, as for example when a trend is present, the moving-range estimate may produce narrower control limits, which will tend to lead to earlier corrective action.

Procedure 18.1 (*X* chart). Determine the central line, control limits, and warning limits for an *X* chart based on a series of *n* independent measurements, which produce the measured values $X_1, X_2, \ldots, X_n$, during a period when the measurement process is in a state of statistical control. At least 2 measurements *must* be used. Ideally, at least 20 measurements should be used.

Procedure:

1. Calculate the sum $\Sigma_{i=1}^{n} X_i$.
2. Calculate the arithmetic mean $\overline{X}$ using the formula

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

3. Calculate an unbiased estimate $\bar{\sigma}$ of the standard deviation (e.g., s / c_4 or $\overline{\mathrm{MR}} / 1.128$).
4. Define the central line, control limits, and warning limits as follows:

$$\mathrm{CL} = \overline{X} \qquad \begin{array}{ll} \mathrm{UCL} = \overline{X} + 3\bar{\sigma} & \mathrm{LWL} = \overline{X} - 2\bar{\sigma} \\ \mathrm{LCL} = \overline{X} - 3\bar{\sigma} & \mathrm{UWL} = \overline{X} + 2\bar{\sigma} \end{array}$$

If n is less than 20, a higher rate of false warnings and failures may occur because of the increased uncertainties of the estimates $\overline{X}$ and $\bar{\sigma}$. So, fewer than 20 measured values should be used only if 20 values cannot be obtained; and the limits should be recalculated when 20 values become available.

EXAMPLE

Problem: Suppose a series of 20 observations of a parameter yield the following normally distributed values.

1,118.9	1,110.5	1,118.3	1,091.0	1,099.8	1,113.7	1,114.4	1,075.1	1,112.8	1,103.7
1,120.5	1,104.0	1,125.7	1,117.6	1,097.6	1,099.8	1,102.3	1,119.9	1,107.8	1,114.9

Determine the central line and warning and control limits for future measurements.

Solution:

Step 1 Calculate $\Sigma X_i = 22{,}168.3$.

Step 2 Calculate the mean $\overline{X} = 22{,}168.3 / 20 = 1{,}108.415$

Step 3 Calculate the experimental standard deviation

$$s = \sqrt{\frac{1}{20 - 1}\sum_{i=1}^{20}(X_i - 1108.415)^2} = 12.044$$

which is based on $\nu = 19$ degrees of freedom. Find $c_4 = 0.98693$ for $\nu = 19$ in Table 18.1 (or estimate $c_4 \approx \frac{4n - 4}{4n - 3} = \frac{76}{77} = 0.9870$), and calculate

$$\bar{\sigma} = \frac{s}{c_4} = \frac{12.044}{0.98693} = 12.2037$$

Step 4 Define the central line, control limits, and warning limits as follows:

$$\begin{aligned} \text{CL} &= 1{,}108.415 \\ \text{UCL} &= 1{,}108.415 + 3(12.2037) = 1{,}145.0 \\ \text{LCL} &= 1{,}108.415 - 3(12.2037) = 1{,}071.8 \\ \text{UWL} &= 1{,}108.415 + 2(12.2037) = 1{,}132.8 \\ \text{LWL} &= 1{,}108.415 - 2(12.2037) = 1{,}084.0 \end{aligned}$$

18A.3 $\overline{X}$ Charts

When subgroup averages are plotted on a control chart, Steps 1 and 2 of Procedure 18.1 may be used to determine the arithmetic mean $\overline{X}$ and the standard deviation $\bar{\sigma}$ of a prior set of data $X_1, X_2, \ldots, X_n$. If k denotes the size of the subgroup, the central line, control limits, and warning limits for the subgroup average are calculated using the formulas

$$CL_{\overline{X}} = \overline{X} \qquad UCL_{\overline{X}} = \overline{X} + 3\bar{\sigma}/\sqrt{k} \qquad UWL_{\overline{X}} = \overline{X} + 2\bar{\sigma}/\sqrt{k}$$
$$LCL_{\overline{X}} = \overline{X} - 3\bar{\sigma}/\sqrt{k} \qquad LWL_{\overline{X}} = \overline{X} - 2\bar{\sigma}/\sqrt{k}$$

If n is less than about 20, a higher rate of false warnings and failures may occur because of the increased uncertainties of the estimates $\overline{X}$ and $\bar{\sigma}$. For this reason fewer than 20 measured values should be used only if 20 values cannot be obtained.

EXAMPLE

Problem: Use the data from the preceding example to determine warning and control limits for subgroup averages when the subgroup size is $k = 5$.

Solution:

Step 1 Calculate $\Sigma X_i = 22{,}168.3$.

Step 2 Calculate the mean $\overline{X} = 22{,}168.3 / 20 = 1{,}108.415$

Step 3 Calculate the experimental standard deviation

$$s = \sqrt{\frac{1}{20-1}\sum_{i=1}^{20}(X_i - 1108.415)^2} = 12.044$$

which is based on $\nu = 19$ degrees of freedom. Find $c_4 = 0.98693$ for $\nu = 19$ in Table 18.1 (or estimate $c_4 \approx \frac{4n-4}{4n-3} = \frac{76}{77} = 0.9870$), and calculate

$$\bar{\sigma} = \frac{s}{c_4} = \frac{12.044}{0.98693} = 12.2037$$

Step 4 Define the central line, control limits, and warning limits as follows:

$$CL_{\overline{X}} = 1{,}108.415$$
$$LCL_{\overline{X}} = 1{,}108.415 - 3(12.2037)/\sqrt{5} = 1{,}092.0$$
$$UCL_{\overline{X}} = 1{,}108.415 + 3(12.2037)/\sqrt{5} = 1{,}124.8$$
$$LWL_{\overline{X}} = 1{,}108.415 - 2(12.2037)/\sqrt{5} = 1{,}097.5$$
$$UWL_{\overline{X}} = 1{,}108.415 + 2(12.2037)/\sqrt{5} = 1{,}119.3$$

18A.4 *R* Charts

The range of a set of values is the difference between the largest value and the smallest. Plotting ranges on a range chart or *R chart* is used to monitor within group variability because *R charts* detect changes in variability more easily. Duplicate measurements for any radiochemistry indicator are made and the difference between the duplicates are used to construct the central line (the mean range), and the control and warning limits in a similar fashion as in the *X chart*. Procedure 18.2 may be used to determine the parameters of the *R chart.*

Procedure 18.2 (*R* chart). Determine the central line and control limits for a *R* chart based on a series of *n* independent sets of duplicate measurements, which produce the values $R_1, R_2, \ldots, R_n$, during a period when the measurement process is in a state of statistical control.

Procedure:

1. Calculate the range, R_i, of each pair of duplicate measurements, (x_i, y_i)

$$R_i = |x_i - y_i|$$

2. Calculate the mean range, $\overline{R}$, using the formula

$$\overline{R} = \frac{1}{n}\sum_{i=1}^{n} R_i$$

3. Calculate the upper control limit as UCL = 3.267 $\overline{R}$.

This approach may also be used for the moving range of a series of individual results.

The factor 3.267 is called "D_4" in references on statistical quality control. The value of D_4 is smaller when the range of a larger group is monitored. When the group size is at least seven, there is also a factor called D_3, which may be used to calculate a lower control limit for the range. Values for D_3 and D_4 are tabulated in *Manual on Presentation of Data and Control Chart Analysis* (ASTM MNL7), as well as many other references.

EXAMPLE

Problem: Suppose a series of 20 duplicate observations of a parameter yield the following pairs of values.

(0.501, 0.491)	(0.490, 0.490)	(0.479, 0.482)	(0.520, 0.512)	(0.500, 0.490)
(0.510, 0.488)	(0.505, 0.500)	(0.475, 0.493)	(0.500, 0.515)	(0.498, 0.501)
(0.523, 0.516)	(0.500, 0.512)	(0.513, 0.503)	(0.512, 0.497)	(0.502, 0.500)
(0.506, 0.508)	(0.485, 0.503)	(0.484, 0.487)	(0.512, 0.495)	(0.509, 0.500)

Determine the central line and upper control limit for the range of future pairs of measurements.

Solution:

Step 1 Calculate the range of each of the 20 pairs .

0.010	0.000	0.003	0.008	0.010
0.022	0.005	0.018	0.015	0.003
0.007	0.012	0.010	0.015	0.002
0.002	0.018	0.003	0.017	0.009

Step 2 Calculate the mean range $\overline{R} = \frac{1}{20}\sum_{i=1}^{20} R_i = \frac{0.189}{20} = 0.00945$

Step 3 Calculate the upper control limit: UCL = 3.267 $\overline{R}$ = (3.267)(0.00945) = 0.0309

18A.5 Control Charts for Instrument Response

A radioactive check source should be used to monitor the efficiency of every radiation counting instrument. MARLAP recommends that the activity and count time for the source be chosen to give no more than 1 percent Poisson counting uncertainty (ANSI N42.23). In other words, at

least 10,000 counts should be obtained in each measurement of the source.

There may be cases when placing a high-activity source in a detector is undesirable, and obtaining 10,000 counts is therefore impractical. The instrument response may not have a Poisson distribution. In this case, if the check source is long-lived, an X or $\overline{X}$ chart based on replicate measurements should be set up. For example, an X or $\overline{X}$ chart is the appropriate efficiency chart for a high-purity germanium detector when the area of a specific photopeak is monitored, since the calculated size of the photopeak may have significant sources of uncertainty in addition to counting uncertainty. An X or $\overline{X}$ chart may be used even if the response is truly Poisson, since the Poisson distribution in this case is approximated well by a normal distribution, but slightly better warning and control limits are obtained by using the unique properties of the Poisson distribution.

Standard guidance documents recommend two types of control charts for Poisson data. A "*c* chart" typically is used in industrial quality control to monitor the number of manufacturing defects per item. A "*u* chart" is used to monitor the number of defects per unit "area of opportunity," when the area of opportunity may vary. Thus, the values plotted on a *c* chart are counts and those plotted on a *u* chart are count rates. The same two types of charts may be adapted for monitoring counts and count rates produced by a radioactive check source. When a *u* chart is used, the "area of opportunity" equals the product of the count time and the source decay factor. In radiation laboratories a variant of the *u* chart is more often used when the count time remains fixed but the decay factor changes during the time when the chart is in use.

Before using control limits derived from the Poisson model, one should use Procedure E1, described in Section 18B.2 of Attachment 18B, to confirm experimentally that the Poisson approximation is adequate and that any excess variance is relatively small at the expected count rate. Factors such as source position that may vary during routine QC measurements should be varied to the same degree during the experiment.

Calculation of warning and control limits using the Poisson model requires only a precise measurement of the source at a time when the instrument is operating properly, preferably near the time of calibration. The precision can be improved either by counting the source longer or by averaging several measurements. In principle both approaches should provide equally good estimates of the count rate; however, an advantage of the latter approach is that it can provide the data needed to detect excess variance (using Procedure E1).

Procedures 18.2 and 18.3, listed below, may be used to determine warning and control limits for measurements of a radioactive check source when the total count follows the Poisson model.

Procedure 18.2 should be used only when the expected count in each measurement is the same, for example when the source is long-lived and all count durations are equal. Procedure 18.3, which implements an alternative to the *u* chart, may be used in all other cases.

Procedure 18.2 (Control chart for Poisson efficiency check data with constant mean). A check source is counted n times on an instrument, producing the measured counts $N_1, N_2, \ldots, N_n$. (Ideally, n is at least 20.) Determine control limits and warning limits for future measurements of the source count on the same instrument.

Procedure:

1. Estimate the central line by

$$\mathrm{CL} = \frac{1}{n}\sum_{i=1}^{n} N_i$$

and the standard deviation by

$$s = \sqrt{\mathrm{CL}}$$

NOTE: The estimate s is biased, but the bias is negligible for the large number of counts typically obtained from a check source.

2. Define the control limits and warning limits (in counts) as follows:

$$\mathrm{UCL} = \mathrm{CL} + 3s \qquad \mathrm{UWL} = \mathrm{CL} + 2s$$
$$\mathrm{LCL} = \mathrm{CL} - 3s \qquad \mathrm{LWL} = \mathrm{CL} - 2s$$

If n is less than 20, a higher rate of false warnings and failures may occur because of the uncertainty in the estimate of the mean. So, fewer than 20 measurements should be used only if 20 measured values are not available.

Procedure 18.3 (Control chart for Poisson efficiency check data with variable mean). A check source is counted n times ($n \geq 1$) on an instrument, producing the measured counts $N_1, N_2, \ldots, N_n$. (It is assumed that the background level is negligible when compared to the source count rate.) Let t_i denote the duration of the i^{th} measurement and d_i the decay factor (for example, $\exp(-\lambda(\Delta t + 0.5\,t_i))$). Determine control limits and warning limits for a future measurement of the source count on the same instrument when the counting period is T and the decay factor is D.

Procedure:

1. Compute the sums $\Sigma_{i=1}^{n} N_i$ and $\Sigma_{i=1}^{n} t_i d_i$.
2. Estimate the mean decay-corrected count rate by

$$\hat{r} = \frac{\Sigma_{i=1}^{n} N_i}{\Sigma_{i=1}^{n} t_i d_i}$$

3. Estimate the central line by

$$\text{CL} = \hat{r} TD$$

and the standard deviation s by

$$s = \sqrt{\text{CL}}$$

4. Define the control limits and warning limits as follows:

$$\text{UCL} = \text{CL} + 3s \qquad \text{UWL} = \text{CL} + 2s$$
$$\text{LCL} = \text{CL} - 3s \qquad \text{LWL} = \text{CL} - 2s$$

If $\Sigma t_i d_i < 20\,TD$, a higher rate of false warnings and failures may occur because of increased uncertainty in the estimate of the count rate $\hat{r}$.

EXAMPLE

Problem: A source containing ^{90}Sr and ^{90}Y in equilibrium is used for efficiency checks on a proportional counter. Near the time of calibration, a series of twenty 600-s measurements are made. The observed counts are as follows:

12,262	12,561	12,606	12,381	12,394	12,518	12,399	12,556	12,565	12,444
12,432	12,723	12,514	12,389	12,383	12,492	12,521	12,619	12,397	12,562

Assume all twenty measurements are made approximately at time 0, so the ten decay factors d_i are all equal to 1. Use Procedure 18.3 to calculate lower and upper control limits for a 600-s measurement of the same source at a time exactly 1 year later.

Solution:

Step 1 Compute the sums $\Sigma N_i = 249{,}718$ and $\Sigma t_i d_i = 12{,}000$.

Step 2 Calculate $\hat{r} = \dfrac{\Sigma N_i}{\Sigma t_i d_i} = \dfrac{249{,}718}{12{,}000} = 20.80983$.

Step 3 The decay time for the final measurement is 1 y = 31,557,600 s. The corresponding decay factor is $D = 0.976055$. The count time is $T = 600$ s. So, compute

$$\text{CL} = (20.80983)(600)(0.976055) = 12{,}187$$

and

$$s = \sqrt{12{,}187} = 110.39$$

Step 4 The control limits and warning limits are

$$\begin{aligned} \text{UCL} &= 12{,}187 + 3 \times 110.39 = 12{,}518 \\ \text{LCL} &= 12{,}187 - 3 \times 110.39 = 11{,}856 \\ \text{UWL} &= 12{,}187 + 2 \times 110.39 = 12{,}408 \\ \text{LWL} &= 12{,}187 - 2 \times 110.39 = 11{,}966 \end{aligned}$$

If substantial excess (non-Poisson) variance is present in the data, the simple Poisson charts described above should not be used. The *c* chart may be replaced by an *X* chart or $\overline{X}$ chart, but a new type of chart is needed to replace the *u* chart. To determine warning and control limits for this chart, one must determine the relative excess variance of the data ξ^2. A value of ξ^2 may be assumed or it may be estimated using procedures described in Attachment 18B. Then Procedure 18.3 may be replaced by the Procedure 18.4, shown below.

Procedure 18.4 (Control chart for Poisson efficiency check data with excess variance). A check source is counted n times on an instrument, producing the measured counts $N_1, N_2, \ldots, N_n$. Let t_i denote the duration of the i^{th} measurement and d_i the decay factor. Let the data follow an approximately Poisson distribution with relative excess variance ξ^2. Determine control limits and warning limits for a future measurement of the source count on the same instrument when the counting period is T and the decay factor is D.

Procedure:

1. Compute the sums $\Sigma_{i=1}^{n} N_i$ and $\Sigma_{i=1}^{n} t_i d_i$.
2. Estimate the mean decay-corrected count rate $\hat{r}$ by

$$\hat{r} = \frac{\sum_{i=1}^{n} \frac{N_i}{1 + r_0 t_i d_i \xi^2}}{\sum_{i=1}^{n} \frac{1}{1 + r_0 t_i d_i \xi^2}} \qquad \text{where} \qquad r_0 = \frac{\sum_{i=1}^{n} N_i}{\sum_{i=1}^{n} t_i d_i}$$

3. Estimate the central line by

$$\mathrm{CL} = \hat{r} TD$$

and the standard deviation s by

$$s = \sqrt{\mathrm{CL} + \xi^2 \mathrm{CL}^2}$$

4. Define the control limits and warning limits as follows:

$$\mathrm{UCL} = \mathrm{CL} + 3s \qquad \mathrm{UWL} = \mathrm{CL} + 2s$$
$$\mathrm{LCL} = \mathrm{CL} - 3s \qquad \mathrm{LWL} = \mathrm{CL} - 2s$$

18A.6 References

American National Standard Institute (ANSI) N42.23. *Measurement and Associated Instrumentation Quality Assurance for Radioassay Laboratories.* 1996.

American Society for Testing and Materials (ASTM) D6299, *Standard Practice for Applying Statistical Quality Assurance Techniques to Evaluate Analytical Measurement System Performance,* 2000

American Society for Testing and Materials (ASTM) E882, Standard Guide for *Accountability and Quality Control in the Chemical Analysis Laboratory.*

American Society for Testing and Materials (ASTM) MNL 7, *Manual on Presentation of Data and Control Chart Analysis* ASTM Manual Series, 6th Edition, 1990.

National Bureau of Standards (NBS). 1964. *Handbook of Mathematical Functions*. M. Abramowitz and Stegun, I., Editors.

Attachment 18B: Statistical Tests for QC Results

18B.1 Introduction

Attachment 18A describes several types of control charts that may be used for statistical quality control in the laboratory. This attachment describes additional statistical methods that may be used, where appropriate, to test the performance of measurement results from blank, replicate, LCS, spikes, CRM, yield-monitor, background, efficiency, calibration, or peak resolution results, with special emphasis on instrumentation results.

18B.2 Tests for Excess Variance in the Instrument Response

As noted in Chapter 19, the counting uncertainty given by the Poisson approximation does not describe the total variability in a counting measurement. A number of factors may generate a small excess component of variance. When a large number of counts are obtained in the measurement, the relative magnitude of the Poisson variance is small; so, the excess component may dominate.

Regardless of whether replication or the Poisson approximation is used to estimate counting uncertainties, MARLAP recommends that a series of check source measurements be made on each instrument periodically to test for excess variance. Procedure E1, which is presented below, may be used to evaluate the measurement results. To check the stability of the instrument itself, one should perform the measurements while holding constant any controllable factors, such as source position, that might increase the variance. To check the variance when such factors are not constant, one may use Procedure E1 but vary the factors randomly for each measurement.

Assume n measurements of the source produce the counts $N_1, N_2, \ldots, N_n$. If the expected count for each measurement is at least 20, so that the Poisson distribution is approximated by a normal distribution, and if the average decay-corrected count rate $\hat{r}$ is determined with adequate precision, then the quantity

$$\chi^2 = \frac{1}{\hat{r}} \sum_{i=1}^{n} \left(\frac{N_i}{t_i d_i} - \hat{r} \right)^2 t_i d_i \qquad (1)$$

where t_i and d_i are the count time and source decay factor for the i^{th} measurement, respectively,

should be distributed approximately as chi-square with $n - 1$ degrees of freedom.[5] The precision of the estimate $\hat{r}$ should be adequate for the test as long as the expected count for each measurement is at least 20. Since a check source is involved, the expected count is usually much greater than 20.

Procedure E1. Determine whether a series of measurements of a check source provide evidence of variance in excess of the Poisson counting variance. Let N_i denote the count observed in the i^{th} measurement. Let $w_i = t_i d_i$, where t_i denotes the count time and d_i denotes the source decay factor (if relevant). If all the values w_i are equal, one may use $w_i = 1$ instead for all i. It is assumed either that the background count rate is negligible or that the decay factors are all nearly equal, so that the expected count in each measurement is proportional to w_i.[6] The procedure tests the null hypothesis that the total measurement variance is the Poisson counting variance.

Procedure:

1. Choose the significance level α.
2. Calculate the sums $\sum_{i=1}^{n} N_i$ and $\sum_{i=1}^{n} w_i$.
3. Estimate the mean decay-corrected count rate by

$$\hat{r} = \frac{\sum_{i=1}^{n} N_i}{\sum_{i=1}^{n} w_i} \tag{2}$$

4. Calculate the chi-square statistic as follows:

$$\chi^2 = \frac{1}{\hat{r}} \sum_{i=1}^{n} \left(\frac{N_i}{w_i} - \hat{r} \right)^2 w_i \tag{3}$$

5. Determine the quantile $\chi^2_{1-\alpha}(n - 1)$ (see Table G.1 in Appendix G). Reject the null

---

[5] If r denotes the true mean decay-corrected count rate, then under the null hypothesis each measured count rate $N_i / t_i d_i$ is approximately normal with mean r and variance $r / t_i d_i$, and the least-squares estimator for r is $\hat{r} = \sum N_i / \sum t_i d_i$. So, the sum $\sum (N_i / t_i d_i - \hat{r})^2 / (r / t_i d_i)$ is approximately chi-square with $n - 1$ degrees of freedom. If $\hat{r}$ is determined accurately, the true mean count rate r may be replaced in the formula by its estimated value $\hat{r}$ to obtain the formula that appears in the text. If all the products $t_i d_i$ are equal, they cancel out of the sum, which becomes $\sum (N_i - \overline{N})^2 / \overline{N}$, as described by Evans (1955), Goldin (1984), and Knoll (1989).

[6] The expected gross count for the i^{th} measurement equals $R_B t_i + r w_i$, where r is the mean net count rate at time 0. The expected count is proportional to w_i if $R_B = 0$, or if all the decay factors are equal so that $t_i \propto w_i$.

hypothesis if and only if the calculated value of χ^2 is greater than $\chi^2_{1-\alpha}(n-1)$. In this case conclude that the variance is greater than predicted by the Poisson model.

EXAMPLE

Problem: A long-lived source is counted $n = 20$ times in a gross radiation detector and the duration of each measurement is 300 s. The following total counts are measured:

11,189	11,105	11,183	10,910	10,998	11,137	11,144	10,751	11,128	11,037
11,205	11,040	11,257	11,176	10,976	10,998	11,023	11,199	11,078	11,149

Are these data consistent with the assumption that the measurement variance is no greater than predicted by the Poisson model? Use 5 percent as the significance level.

Solution:

Step 1 The significance level is specified to be $\alpha = 0.05$.

Step 2 Since the source is long-lived and all the count times are equal, let $w_i = 1$ for each i. Calculate $\Sigma N_i = 221{,}683$ and $\Sigma w_i = 20$.

Step 3 Calculate the mean count rate $\hat{r} = 221{,}683 / 20 = 11{,}084.15$.

Step 4 Calculate the chi-square statistic

$$\chi^2 = \frac{1}{\hat{r}} \sum_{i=1}^{n} \left(\frac{N_i}{w_i} - \hat{r} \right)^2 w_i = \frac{1}{11{,}084.15} \sum_{i=1}^{20} (N_i - 11{,}084.15)^2 = 24.87$$

Step 5 The number of degrees of freedom is $20 - 1 = 19$. According to Table G.1, the 0.95-quantile for a chi-square distribution with 19 degrees of freedom is 30.14. Since $24.87 \leq 30.14$, do not reject the null hypothesis. The data are consistent with the assumption of Poisson counting statistics at the 5 percent significance level.

A two-sided version of Procedure E1 may also be used to test whether the measurement variance is either greater than or less than predicted by the Poisson model. Step 5 must be changed so that the null hypothesis is rejected if the value of the test statistic χ^2 does not lie between the two quantiles $\chi^2_{\alpha/2}(n-1)$ and $\chi^2_{1-\alpha/2}(n-1)$.

A chi-square test may require many measurements or long count times to detect a small excess variance component. When all measurements have the same expected count μ, the detection limit for the *relative* excess variance, or its minimum detectable value, is equal to

$$\xi_D^2 = \frac{1}{\mu}\left(\frac{\chi^2_{1-\alpha}(n-1)}{\chi^2_{\beta}(n-1)} - 1\right) \tag{4}$$

where β is the specified probability of a type II error (failure to detect) (Currie 1972). Note that since ξ_D^2 represents a relative variance, its square root ξ_D represents a relative standard deviation.

EXAMPLE: A long-lived source is counted 20 times, and each measurement has the same duration. The average of the measured counts is 10,816. If $\alpha = \beta = 0.05$, the minimum detectable value of the relative excess variance is estimated by

$$\xi_D^2 = \frac{1}{10{,}816}\left(\frac{\chi^2_{0.95}(19)}{\chi^2_{0.05}(19)} - 1\right) = \frac{1}{10{,}816}\left(\frac{30.14}{10.12} - 1\right) = \frac{1.978}{10{,}816} = 1.829\times10^{-4}$$

which corresponds to a relative standard deviation $\xi_D = \sqrt{1.829\times10^{-4}} = 0.01352$, or about 1.35 percent.

If (1) the relative excess variance in a measurement is not affected by count time, (2) a fixed total count time is available, and (3) all measurements have the same expected count (e.g., when all count times are equal and the source is long-lived), then it is possible to determine the number of measurements that minimizes ξ_D^2 (Currie 1972). The optimal number is the number *n* that minimizes the quantity

$$F(n) = n\left(\frac{\chi^2_{1-\alpha}(n-1)}{\chi^2_{\beta}(n-1)} - 1\right) \tag{5}$$

The solution may be found by computing $F(n)$ for $n = 2, 3, 4, \ldots$, until the computed value begins to increase. When $\alpha = \beta = 0.05$, the optimal number of measurements is $n = 15$, although the improvement as *n* increases from 6 to 15 is slight. If *n* is increased further, the detection limit ξ_D^2 worsens unless the total count time is also increased.

A chi-square test may also be used to test whether the total source measurement variance consists

of a Poisson component and a specified excess component (Currie 1972). Procedure E2, described below, implements this test. If the specified component is zero, Procedure E2 is equivalent to E1.

Procedure E2. Determine whether a series of measurements of a check source provide evidence that the measurement variance is greater than the Poisson component plus a specified excess component. (Refer to the notation used in Procedure E1.) Let ξ^2 denote the value of the relative excess variance under the null hypothesis H_0.

Procedure:

1. Choose the significance level α.

2. Calculate the sums $\sum_{i=1}^{n} N_i$ and $\sum_{i=1}^{n} w_i$, where $N_1, N_2, \ldots, N_n$ are the measured values.

3. Estimate the mean decay-corrected count rate $\hat{r}$ in two steps by

$$r_0 = \frac{\sum_{i=1}^{n} N_i}{\sum_{i=1}^{n} w_i} \qquad \text{and} \qquad \hat{r} = \sum_{i=1}^{n} \frac{N_i}{1 + r_0 w_i \xi^2} \Big/ \sum_{i=1}^{n} \frac{w_i}{1 + r_0 w_i \xi^2} \tag{6}$$

   (If $w_1 = w_2 = \cdots = w_n$ or $\xi^2 = 0$, then $\hat{r} = r_0$.)

4. Calculate the chi-square statistic as follows:[7]

$$\chi^2 = \sum_{i=1}^{n} \frac{(N_i / w_i - \hat{r})^2}{\hat{r} / w_i + \hat{r}^2 \xi^2} \tag{7}$$

5. Determine the quantile $\chi^2_{1-\alpha}(n-1)$ (see Table G.1). Reject the null hypothesis if and only if the calculated value of χ^2 is greater than $\chi^2_{1-\alpha}(n-1)$. In this case conclude that the relative excess variance is greater than ξ^2.

Procedure E2, like E1, can easily be converted to a two-sided test by changing Step 5.

[7] In Currie (1972), the variance of N_i is estimated by $N_i + \xi^2 N_i^2$. The estimated variance used here is calculated by pooling the counting data to reduce any small bias caused by the correlation between N_i and $N_i + \xi^2 N_i^2$.

The excess component may be estimated by solving Equations 18.6 and 18.7 for the value of ξ that gives $\chi^2 = n - 1$. An iterative computer algorithm, such as bisection, which repeatedly tries values of ξ and computes χ^2 can be used.[8] An approximate confidence interval for the relative excess variance may similarly be found by solving for values of ξ which give $\chi^2 = \chi^2_{(1\pm\gamma)/2}(n-1)$, where γ is the desired confidence coefficient (Currie, 1972).

If $w_1 = w_2 = \cdots = w_n$, the iterative algorithm is unnecessary. In this case the value of ξ may be estimated directly using the formula

$$\xi^2 = \frac{1}{\overline{N}^2}\left(\frac{1}{n-1}\sum_{i=1}^{n}(N_i - \overline{N})^2 - \overline{N}\right) \tag{8}$$

or by $\xi = 0$ if the preceding formula gives a negative result. Similarly, the approximate lower confidence limit is given by the formula

$$\xi^2_{\text{lower}} = \frac{1}{\overline{N}^2}\left(\frac{1}{\chi^2_{(1+\gamma)/2}(n-1)}\sum_{i=1}^{n}(N_i - \overline{N})^2 - \overline{N}\right) \tag{9}$$

and the approximate upper confidence limit is given by

$$\xi^2_{\text{upper}} = \frac{1}{\overline{N}^2}\left(\frac{1}{\chi^2_{(1-\gamma)/2}(n-1)}\sum_{i=1}^{n}(N_i - \overline{N})^2 - \overline{N}\right) \tag{10}$$

EXAMPLE

Problem: A long-lived efficiency check source is counted once a day for 20 days, and each measurement has the same duration. Suppose the measured counts (N_i) are:

14,454	15,140	15,242	14,728	14,756	15,040	14,768	15,128	15,150	14,872
14,845	15,511	15,032	14,746	14,731	14,982	15,047	15,272	14,765	15,143

[8] Newton's method, which converges more rapidly, can also be used, but its use is more practical if one replaces $\hat{r}$ by r_0 in the denominator of each term of Equation 18.7.

Use these data to estimate ξ and determine a 95 percent two-sided confidence interval for its value.

Solution: Since the source is long-lived and all the measurements have the same duration, $w_1 = w_2 = \cdots = w_{20}$ and Equations 18.8 through 18.10 may be used. So, calculate $\Sigma N_i = 299{,}352$ and $\overline{N} = 299{,}352 / 20 = 14{,}967.6$. Then the value of ξ is estimated as

$$\xi = \frac{1}{14{,}967.6}\sqrt{\frac{1}{20-1}\sum_{i=1}^{20}(N_i - 14{,}967.6)^2 - 14{,}967.6} = 0.014463$$

The 95 percent confidence limits are calculated as follows:

$$\xi_{\text{lower}} = \frac{1}{\overline{N}}\sqrt{\frac{1}{\chi^2_{0.975}(20-1)}\sum_{i=1}^{20}(N_i - \overline{N})^2 - \overline{N}}$$

$$= \frac{1}{14{,}967.6}\sqrt{\frac{1}{32.852}\sum_{i=1}^{20}(N_i - 14{,}967.6)^2 - 14{,}967.6}$$

$$= 0.0096334$$

$$\xi_{\text{upper}} = \frac{1}{\overline{N}}\sqrt{\frac{1}{\chi^2_{0.025}(20-1)}\sum_{i=1}^{20}(N_i - \overline{N})^2 - \overline{N}}$$

$$= \frac{1}{14{,}967.6}\sqrt{\frac{1}{8.9065}\sum_{i=1}^{20}(N_i - 14{,}967.6)^2 - 14{,}967.6}$$

$$= 0.022846$$

For most practical purposes the excess variance may be considered negligible in a counting measurement if the total count N is less than $1 / 10\xi^2$, since, in this case, the excess variance increases the standard deviation of the measured count by less then 5 percent. Similarly, the counting variance may be considered negligible if $N \geq 10 / \xi^2$.

EXAMPLE: Suppose $N = 1{,}000$ counts observed in a measurement and ξ has been estimated to be 0.01. Then $N = 1 / 10\xi^2$. The standard uncertainty of N is evaluated as

$$u(N) = \sqrt{N + \xi^2 N^2} = \sqrt{1{,}000 + 10^{-4}10^6} = \sqrt{1{,}100} \approx 1.05\sqrt{N}$$

If $N = 100{,}000$, then $N = 10 / \xi^2$ and

$$u(N) = \sqrt{10^5 + 10^{-4}10^{10}} = \sqrt{1{,}100{,}000} \approx 1.05(\xi N)$$

So, $u(N) \approx \sqrt{N}$ for $N \leq 1{,}000$, and $u(N) \approx \xi N$ for $N \geq 100{,}000$.

18B.3 Instrument Background Measurements

This section presents statistical tests related to measurements of instrument background levels. The tests are intended for single-channel detectors but may be applied to multichannel systems if wide spectral regions are integrated. Tests are described for comparing background levels to preset limits, for detecting changes in background levels between measurements, and for detecting the presence of variability in excess of that predicted by the Poisson model.

18B.3.1 Detection of Background Variability

The chi-square test (Procedure E1) used to detect excess variance in measurements of a check source may be adapted for background measurements. Procedure B1 implements a chi-square test for backgrounds. This test is one-sided, although Step 6 can be modified to implement a two-sided test.

Procedure B1. Determine whether a series of measurements of an instrument's background provide evidence of variance in excess of the Poisson counting variance. Let N_i denote the count observed in the i^{th} measurement, and let t_i denote the count time.

Procedure:

1. Determine the significance level α.
2. Calculate the sums $\Sigma_{i=1}^{n} N_i$ and $\Sigma_{i=1}^{n} t_i$.

3. Estimate the mean background count rate by

$$\hat{r} = \frac{\sum_{i=1}^{n} N_i}{\sum_{i=1}^{n} t_i} \tag{11}$$

4. Let $t_{\min}$ be the smallest value of t_i. If $\hat{r}t_{\min} \geq 20$, go to Step 5. Otherwise, discard all measured values N_i for which $\hat{r}t_i < 20$. If possible, restart the test at Step 2; if not, stop.

5. Calculate the chi-square statistic as follows:

$$\chi^2 = \frac{1}{\hat{r}} \sum_{i=1}^{n} \left(\frac{N_i}{t_i} - \hat{r} \right)^2 t_i \tag{12}$$

6. Determine the quantile $\chi^2_{1-\alpha}(n-1)$ (see Table G.1 in Appendix G). Reject the null hypothesis if and only if the calculated value of χ^2 is greater than $\chi^2_{1-\alpha}(n-1)$. In this case, conclude that the instrument background does not follow the Poisson model.

EXAMPLE

Problem: Twenty overnight background measurements are performed on a proportional counter. The duration of each measurement is 60,000 s, and the following alpha counts are measured:

14	23	23	25	28	22	19	26	20	27
30	21	34	32	24	27	25	19	19	25

Are these data consistent with the assumption that the measurement variance is attributable to Poisson counting statistics? Use 5 percent as the significance level.

Solution:

Step 1 The significance level is specified to be $\alpha = 0.05$.

Step 2 Calculate $\Sigma N_i = 483$ and $\Sigma t_i = 20 \times 60{,}000 = 1{,}200{,}000$.

Step 3 Calculate the mean count rate $\hat{r} = 483/1{,}200{,}000 = 0.0004025$.

Step 4 Since $t_{\min} = 60{,}000$, $\hat{r}t_{\min} = 24.15$. Since $24.15 \geq 20$, go to Step 5.

Step 5 Calculate the chi-square statistic

$$\chi^2 = \frac{1}{\hat{r}} \sum_{i=1}^{n} \left(\frac{N_i}{t_i} - \hat{r} \right)^2 t_i = \frac{1}{0.0004025} \sum_{i=1}^{20} \left(\frac{N_i}{60{,}000} - 0.0004025 \right)^2 60{,}000 = 18.49$$

Step 6 The number of degrees of freedom is 20 – 1 = 19. According to Table G.1, the 0.95-quantile for a chi-square distribution with 19 degrees of freedom is 30.14. Since 18.49 ≤ 30.14, do not reject the null hypothesis. The data are consistent with the Poisson model.

All the background tests described below are based on the assumption of Poisson counting statistics. If Procedure B1 indicates the Poisson assumption is invalid, each test requires modification or replacement. In most cases, unless the observed background counts are very low, standard statistical tests for normally distributed data may be used instead (e.g., NBS, 1963; EPA, 1998).

18B.3.2 Comparing a Single Observation to Preset Limits

High background levels on an instrument degrade detection capabilities and may indicate the presence of contamination. Unusually low levels on certain types of instruments may indicate instrument failure. When these issues are of concern, one or both of the two statistical tests described below may be performed to determine whether the true background level is outside of its desired range.

The result of the background measurement in counts is assumed to have a Poisson distribution. In both of the following tests, t denotes the count time, and r denotes the preset lower or upper limit for the true mean background count rate R_B. Given an observed count N_B, Procedure B2 determines whether $R_B > r$ and B3 determines whether $R_B < r$.

Procedure B2 should be used when r is an upper limit and B3 should be used when r is a lower limit. Thus, the background level is assumed to be within its acceptable limits unless there is statistical evidence to the contrary. The alternative approach, which changes the burden of proof, may be used if rt is large enough.

If rt is extremely large (e.g., if $rt \geq 2{,}500$), there is probably no justification for a statistical test. Instead, the observed count rate may be compared directly to r.

Procedure B2. Determine whether the mean background count rate R_B is greater than r. Test the null hypothesis H_0: $R_B \leq r$ against the alternative hypothesis H_1: $R_B > r$.

Procedure:

1. Choose the significance level α.

2. If $N_B \leq rt$, conclude that there is insufficient evidence to reject the null hypothesis, and stop. Otherwise, if $rt < 20$, go to Step 6. If $rt \geq 20$, go to Step 3.

3. Calculate

$$Z = \frac{0.5 + N_B - rt}{\sqrt{rt}} \tag{14}$$

4. Determine $z_{1-\alpha}$, the $(1-\alpha)$-quantile of the standard normal distribution (see Table G.1 in Appendix G).

5. Reject the null hypothesis if and only if $Z > z_{1-\alpha}$. Stop.

   NOTE: If the background count time t is always the same, a fixed upper control limit may be calculated using the formula

$$\text{UCL} = \textbf{round}\left(rt + z_{1-\alpha}\sqrt{rt}\right)$$

   where **round** denotes the function that rounds its argument to the nearest integer. Then Steps 3–5 are effectively performed by comparing the observed value N_B to UCL.

6. Determine $\chi^2_\alpha(2N_B)$, the α-quantile of the chi-square distribution with 2NB degrees of freedom (see Table G.1 in Appendix G), and calculate $Q = 0.5\,\chi^2_\alpha(2N_B)$.

7. Reject the null hypothesis if and only if $Q > rt$.

EXAMPLE

Problem: To ensure adequate detection capabilities, a laboratory establishes an upper limit of 0.02 cps for beta backgrounds on a proportional counter. A 6,000-s background measurement is performed, during which 125 beta counts are observed. Determine whether this measurement result gives 95 percent confidence that the background is greater than 0.02 cps.

Solution:	The values of the variables are $N_B = 125$, $t = 6{,}000$ and $r = 0.02$.
Step 1	The significance level α is $1 - 0.95 = 0.05$.
Step 2	Since $N_B \geq rt = 120$ and $rt \geq 20$, go to Step 3.
Step 3	Calculate $Z = (0.5 + 125 - 120) / \sqrt{120} = 0.5021$.
Step 4	Table G.1 shows that $z_{0.95} = 1.645$.
Step 5	Since $0.5021 \leq 1.645$, do not reject the null hypothesis. There is insufficient evidence to conclude that the beta background exceeds 0.02 cps.

EXAMPLE

Problem: The same laboratory establishes an upper limit of 0.002 cps for alpha backgrounds on the same counter. A 6,000-s background measurement is performed, during which 19 alpha counts are observed. Determine whether this measurement result gives 95 percent confidence that the background is greater than 0.002 cps.

Solution:	The values of the variables are $N_B = 19$, $t = 6{,}000$ and $r = 0.002$.
Step 1	The significance level α is $1 - 0.95 = 0.05$.
Step 2	Since $N_B \geq rt = 12$ and $rt < 20$, go to Step 6.
Step 6	Table G.1 shows that $\chi^2_{0.05}(38) = 24.88$. So, $Q = 0.5 \cdot 24.88 = 12.44$.
Step 7	Since $12.44 > 12$, reject the null hypothesis. The data give 95 percent confidence that the alpha background is greater than 0.002 cps.

Procedure B3. Determine whether the mean background count rate R_B is less than r. Test the null hypothesis H_0: $R_B \geq r$ against the alternative hypothesis H_1: $R_B < r$.

Procedure:

1. Choose the significance level α.

2. If $N_B \geq rt$, conclude that there is insufficient evidence to reject the null hypothesis, and stop. Otherwise, if $rt < 20$, go to Step 6. If $rt \geq 20$, go to Step 3.

3. Calculate

$$Z = \frac{0.5 + N_B - rt}{\sqrt{rt}} \tag{15}$$

4. Determine $z_{1-\alpha}$, the $(1-\alpha)$-quantile of the standard normal distribution (see Table G.1 in Appendix G).

5. Reject the null hypothesis if and only if $Z < -z_{1-\alpha}$. Stop.

NOTE: If the background count time t is always the same, a lower control limit may be calculated using the formula

$$\text{LCL} = \textbf{round}\left(rt - z_{1-\alpha}\sqrt{rt}\right).$$

Steps 3–5 are then effectively performed by comparing N_B to LCL.

6. Determine $\chi^2_{1-\alpha}(2N_B + 2)$, the $(1-\alpha)$-quantile of the chi-square distribution with $2N_B + 2$ degrees of freedom (see Table G.1), and calculate $Q = 0.5\,\chi^2_{1-\alpha}(2N_B + 2)$.

7. Reject the null hypothesis if and only if $Q < rt$.

EXAMPLE

Problem: A laboratory establishes a lower limit of 0.01 cps for beta backgrounds on a proportional counter. A 6,000-s background measurement is performed, during which 50 beta counts are observed. Determine whether this measurement result gives 95 percent confidence that the background is less than 0.01 cps.

Solution: The values of the variables are $N_B = 50$, $t = 6{,}000$ and $r = 0.01$.

Step 1 The significance level α is $1 - 0.95 = 0.05$.

Step 2 Since $N_B \le rt = 60$ and $rt \ge 20$, go to Step 3.

Step 3 Calculate $Z = (0.5 + 50 - 60) / \sqrt{60} = -1.226$.

Step 4 Table G.1 shows that $z_{0.95} = 1.645$.

Step 5 Since $-1.226 \ge -1.645$, do not reject the null hypothesis.

18B.3.3 Comparing the Results of Consecutive Measurements

If consecutive measurements of the background level on an instrument give significantly different values, one should be concerned about the accuracy of any laboratory sample measurements made between the two background measurements. If the background has increased, the laboratory sample activities may have been overestimated. If the background has decreased, the activities may have been underestimated.

Let N_1 and N_2 denote the counts observed in two independent background measurements on the same instrument, and assume they represent Poisson distributions with unknown means. Let t_1 and t_2 denote the corresponding count times. The following two procedures may be used to determine whether the difference between the two observed values is significantly larger than would be expected on the basis of the Poisson model. Procedure B4 determines whether the second value is significantly greater than the first. Procedure B5 determines whether there is a significant difference between the two values.

Procedure B4. Determine whether the second mean background count rate R_2 is higher than the first R_1. Test the null hypothesis H_0: $R_1 \geq R_2$ against the alternative hypothesis H_1: $R_1 < R_2$.

Procedure:

1. Choose the significance level α.

2. If $N_1 / t_1 \geq N_2 / t_2$, conclude that there is insufficient evidence to reject the null hypothesis, and stop. Otherwise, if $N_1 \geq 20$ and $N_2 \geq 20$, go to Step 3. If $N_1 < 20$ or $N_2 < 20$, go to Step 6.

3. Calculate

$$Z = \left(\frac{N_2}{t_2} - \frac{N_1}{t_1} \right) \Bigg/ \sqrt{\frac{N_1 + N_2}{t_1 t_2}} \qquad (16)$$

4. Determine $z_{1-\alpha}$, the $(1-\alpha)$-quantile of the standard normal distribution.

5. Reject the null hypothesis if and only if $Z > z_{1-\alpha}$. Stop.

6. Let $p = t_1 / (t_1 + t_2)$ and $q = t_2 / (t_1 + t_2)$. If $N_1 < N_2$, calculate

$$S = \sum_{k=0}^{N_1} \binom{N_1 + N_2}{k} p^k q^{N_1 + N_2 - k} \tag{17}$$

If $N_1 \geq N_2$, calculate S more efficiently using the formula

$$S = 1 - \sum_{k=N_1+1}^{N_1+N_2} \binom{N_1 + N_2}{k} p^k q^{N_1 + N_2 - k} \tag{18}$$

7. Reject the null hypothesis if and only if $S \leq \alpha$.

EXAMPLE

Problem: A 60,000-s background measurement is performed on an alpha spectrometer and 15 total counts are observed in a particular region of interest. After a test source is counted, a 6,000-s background measurement is performed and 3 counts are observed. Assuming Poisson counting statistics, is the second measured count rate (0.0005 cps) significantly higher than the first (0.00025 cps) at the 5 percent significance level?

Solution: The variables are $N_1 = 15$, $t_1 = 60{,}000$, $N_2 = 3$, and $t_2 = 6{,}000$.

Step 1 The significance level α is specified to be 0.05.

Step 2 Since $N_1 / t_1 = 0.00025 < 0.0005 = N_2 / t_2$, $N_1 < 20$, and $N_2 < 20$, go to Step 6.

Step 6 $p = \frac{60{,}000}{66{,}000} = \frac{10}{11}$ and $q = \frac{6{,}000}{66{,}000} = \frac{1}{11}$. Since $N_1 \geq N_2$, calculate S using the second formula.

$$S = 1 - \left(\binom{18}{16} \left(\frac{10}{11}\right)^{16} \left(\frac{1}{11}\right)^{2} + \binom{18}{17} \left(\frac{10}{11}\right)^{17} \left(\frac{1}{11}\right)^{1} + \binom{18}{18} \left(\frac{10}{11}\right)^{18} \left(\frac{1}{11}\right)^{0} \right)$$

$$= 1 - 0.7788 = 0.2212 .$$

Step 7 Since $S \geq \alpha$, there is not enough evidence to reject the null hypothesis. The second measured count rate is not significantly higher than the first.

Procedure B5. Determine whether the mean background count rates are different. Test the null hypothesis H_0: $R_1 = R_2$ against the alternative hypothesis H_1: $R_1 \neq R_2$.

Procedure:

1. Choose the significance level α.

2. If $N_1 / t_1 = N_2 / t_2$, conclude that there is insufficient evidence to reject the null hypothesis, and stop. Otherwise, if $N_1 < 20$ or $N_2 < 20$, go to Step 6. If $N_1 \geq 20$ and $N_2 \geq 20$, go to Step 3.

3. Calculate Z using Equation 18.17.

4. Determine $z_{1-\alpha/2}$, the $(1 - \alpha / 2)$-quantile of the standard normal distribution.

5. Reject the null hypothesis if and only if $|Z| > z_{1-\alpha/2}$. Stop.

6. If $N_1 / t_1 < N_2 / t_2$, use Procedure B4 with significance level $\alpha / 2$ to determine whether $R_1 < R_2$. If $N_1 / t_1 > N_2 / t_2$, use Procedure B4 with significance level $\alpha / 2$ and with the observations reversed to determine whether $R_2 < R_1$.

18B.4 Negative Activities

When the measured count rate for a test source is less than that of the corresponding instrument background, giving a negative value for the source activity, Procedure B4 may be used to determine whether the difference between the two count rates is significantly more than should be expected on the basis of the Poisson model and the assumption that the source is a blank. (Let N_1 and t_1 be the source count and counting time and let N_2 and t_2 be the background count and counting time.). If a significant difference is found, it may indicate that the background measurement was biased, the true background is variable or non-Poisson, or the instrument is unstable.

18B.5 References

Currie, Lloyd A. 1972. The Limit of Precision in Nuclear and Analytical Chemistry. *Nuclear Instruments and Methods* 100(3): 387–395.

Environmental Protection Agency (EPA). 1998. *Guidance for Data Quality Assessment: Practical Methods for Data Analysis*. EPA QA/G-9, QA97 Version. EPA/600/R-96/084,

2289 EPA, Quality Assurance Division, Washington, DC.

2290 Evans, Robley D. 1955. *The Atomic Nucleus*. McGraw-Hill, New York, NY.

2291 Goldin, Abraham S. 1984. Evaluation of Internal Control Measurements in Radioassay. *Health*
2292 *Physics* 47(3): 361–374.

2293 Knoll, Glenn F. 1989. *Radiation Detection and Measurement*, 2nd ed. John Wiley and Sons, New
2294 York, NY.

2295 National Bureau of Standards (NBS). 1963. *Experimental Statistics*. NBS Handbook 91, National
2296 Bureau of Standards, Gaithersburg, MD.